1963

An investigation of the effects of transportation lag on relay servo systems

Chiang, Shu-ming

Monterey, California: U.S. Naval Postgraduate School

http://hdl.handle.net/10945/12419

AN INVESTIGATION OF THE EFFECTS OF TRANSPORTATION LAG ON RELAY SERVO SYSTEMS

SHU-MIN CHIANG

AN INVESTIGATION OF THE EFFECTS OF TRANSPORTATION LAG ON RELAY SERVO SYSTEMS

* * * * *

Shu-min Chiang

AN INVESTIGATION OF THE EFFECTS OF TRANSPORTATION LAG ON RELAY SERVO SYSTEMS

by

Shu-min Chiang
Lieutenant, Chinese Navy

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE
IN
ELECTRICAL ENGINEERING

United States Naval Postgraduate School
Monterey, California

1 9 6 3

AN INVESTIGATION OF THE EFFECTS OF TRANSPORTATION LAG ON RELAY SERVO SYSTEMS

by

Shu-min Chiang

This work is accepted as fulfilling
the thesis requirements for the degree of
MASTER OF SCIENCE
IN
ELECTRICAL ENGINEERING
from the
United States Naval Postgraduate School

ABSTRACT

Most physical systems are nonlinear to some extent. One variety of nonlinear system uses a relay or other on-off device as a control element.

Since in the process industries or relay servo systems, one often encounters a type of time delay called transportation lag or dead time. It is highly desirable to investigate the effects of the transportation lag on the process or relay servo system.

This paper is to demonstrate the effects of the transportation lag on the relay servo system and show how to eliminate these effects.

Experimental results indicate that the transportation lag causes a continuous oscillation with a steady state over- and under-shoot, the response can never be settled, and also it leads to a longer rise time. In addition to those, the transportation lag results in a terminal limit cycle. Though the tachometer feedback has been used to compensate the system, this oscillation can never be eliminated completely when some delay exists. Therefore, a five percent overshoot and undershoot was considered acceptable.

The author wishes to express his appreciation for the advice, assistance and encouragement given by Dr. Richard Carl Dorf, Professor of Electrical Engineering, United States Naval Postgraduate School, throughout this investigation. He also owes a debt of gratitude to Dr. John R. Ward, Professor of Electrical Engineering Department, U. S. Naval Postgraduate School, for his assistance during the initial stages of the work.

TABLE OF CONTENTS

Section	Title	Page
1.	Introduction	1
2.	Relay servo system	6
	2.1. General description	6
	2.2. Relay servos	7
	2.3. Phase plane trajectory	10
	2.4. On-off control	18
3.	Description of the system	22
4.	Nonlinear analysis	24
	4.1. General	24
	4.2. The describing function	25
	4.3. Computation of the describing function	26
	4.4. Analysis of the limit cycle	28
	4.4.1. Frequency response method	28
	4.4.2. The root locus method	33
5.	System with dead time	38
	5.1. Transfer function associated with pure dead time	38
	5.2. Frequency response of a system with dead time	38
6.	Transportation lag	45
	6.1. General	45
	6.2. Pade approximation	46
	6.3. Simulation of transportation lag	48
	6.4. Generation of time delay from magnetic tape recorder	51
7.	System simulation	57
	7.1. Analog computer simulation	57
	7.2. Digital computer program	59
	7.3. Comparison of the results obtained from the digital computer and from the analog computer	69
8.	Effects of transportation lag on relay servo system	72
	8.1. General	72
	8.2. The frequency response	72
	8.3. The root locus	74
	8.4. Transient response	75
	8.5. Time delay in the phase plane diagram	81
	8.6. The effect of time delay in steady state oscillation frequency	93
	8.7. The effect of the magnitude of the relay voltage	98
9.	Elimination of the steady state oscillation	101
	9.1. Velocity feedback stabilization	101
	9.2. Optimising the transient response	101
10.	The response of relay servo system to ramp function input	115

Table of Contents

Section	Title	Page
11.	Conclusions	125
Bibliography		127

LIST OF ILLUSTRATIONS

Figure	Title	Page
1-1	System in which dead time is often encountered	3
1-2	System in which dead time is often encountered	3
1-3	The effect of delay caused by the water slug	3
1-4	The reaction curve of the process	3
1-5	Curves showing the effect of transportation lag	5
2-1	Block diagram of the contactor servomechanism	6
2-2	Phase plane trajectory for the second order contactor servomechanism	12
2-3	Switching line for contactor servomechanism	13
2-4	Typical phase plane trajectories	15
2-5	Switching point for optimum step response	17
2-6	Two-position control	20
2-7	Two-position control without differential gap	20
2-8	Two-position control with differential gap	20
3-1	Generalized block diagram of basic system	23
4-1	Characteristic of ideal relay	27
4-2	Describing function of ideal relay	28
4-3	Flow diagram for sinusoidal quantities	28
4-4	Curves for predicting limit cycle amplitude and period in an on-off system	31
4-5	Nyquist plot of equation 4-5	32
4-6	Block diagram of the relay servo system	33
4-7	Construction of the required locus with transportation lag	35
5-1	Bode diagram for a pure dead-time unit and a first-order lag unit with time constant of the latter equal to the dead time of the former	39

Figure	Title	Page
5-2	Dead time element	40
5-3	Block diagram of a simple feedback system with transportation lag	42
5-4	Nyquist diagram for the system in Fig. 5-3	42
5-5	Block diagram of a single loop feedback system	43
5-6	Stability locus in the $G_s G_b$-plane for a system of the type shown in Fig. 5-5(b)	43
5-7	Determination of phase margin in a system with transportation lag	43
6-1	The phenomenon of dead time	45
6-2	Analog computer circuit for fourth order Pade approximation	50
6-3	Phase shift characteristics of the Pade approximation	52
6-4	Curves obtained from Pade approximation	53
6-5	Phase shift characteristics of the Pade approximation (eighth order)	54
6-6	Curves obtained from magnetic tape recorder	56
7-1	The block diagram of the relay servo system	57
7-2	The simulation of the ideal relay	57
7-3	Operational diagram of analog computer set up for the block diagram of the relay servo system of Fig. 7-1	58
7-4	The block diagram the relay servo system	59
7-5	Output, velocity of the system with t_d = .4 sec. comparison the results obtained from digital computer and analog computer	70
7-6	Comparison of the phase trajectory of the system with t_d = .4 second	71
8-1	Bode diagram of equation 8-2	73
8-2	Root locus of 8-4	74
8-3	Root locus of relay servo system with time delay of .5 second	76

Figure	Title	Page
8-4(a)	Response of relay servo system, t_d = 0, .1, .25, .5, .8 and 1.0 respectively, k_t = .02	77
8-4(b)	Response of relay servo system with t_d = 0, .1, .25, .5, .8 and 1.0 respectively, k_t = .8	78
8-5	M_{p_t}, and t_p vs. delay time	80
8-6	Time rise vs. delay time	82
8-7(a)	The plant output, velocity and system error of an undelayed relay servo system	83
8-7(b)	The plant output, velocity and error of relay servo system, t_d = .5 sec. k_t = .02	84
8-7(c)	The plant output, velocity and system error of relay servo system, t_d = 1.0 sec. k_t = .02	85
8-7(d)	The plant output, velocity and error of relay servo system, t_d = 2.5 sec. k_t = .02	86
8-8	Normalized phase diagram for relay servo system with a time delay of .1 second for a step unit	87
8-9	Phase plane trajectory showing loci of limit cycle for oscillation of constant amplitude with various time delay	89
8-10	Curve showing the relation between time delay and amplitude of oscillation	90
8-11(a)	Phase plane diagram for a relay servo system showing the effect of time delay (t_d = 0, .5)	91
8-11(b)	Phase plane diagram for relay servo system showing the effect of time delay (t_d = 1.0, 2.5)	92
8-12	Error vs. time curves showing the effect of time delay on the system error	95
8-13	Velocity vs. time curves showing the effect of time delay	96
8-14	Steady state oscillation frequency vs. time delay	97
8-15	System error against time delay	97
8-16	Curves showing the effect of the relay voltage on the response, t_d = .5, k_t = .02	99
8-17	Phase plane diagram showing the effect of relay voltage, t_d = .5, k_t = .02	100

9-1	Phase plane of relay servo system with a time delay of .25 second compensated by increasing the gain of tachometer feedback to .8	102
9-2	Curve showing the response of the relay servo system be compensated to an acceptable optimum operation by increasing k_t to a value of .8, time delay .25 second	103
9-3(a)	Response and phase plane, t_d = .0, k_t = .3	106
9-3(b)	Response and phase plane, t_d = .1 sec. k_t = .4	107
9-3(c)	Response and phase plane, t_d = .25, k_t = .8	108
9-3(d)	Response and phase plane, t_d = .5, k_t = 1.5	109
9-3(e)	Response and phase plane, t_d = 1.0, k_t = 2.4	110
9-4	Curve showing system with .8 second time delay can not be compensated	111
9-5	Curve showing system with one second time delay can not be compensated	111
9-6	Time delay vs. tachometer gain curve	112
10-1	C, Cdot, E vs. time, ramp input	116
10-2	Phase trajectory of relay servo system, relay voltage .5 volt, ramp input	117
10-3	C, Cdot, E vs. time, ramp input, t_d = 0, k_t = .02, V = 1.5 volt	118
10-4	Cdot vs. C phase plane of a system with t_d = 0, k_t = .02, V = 1.5 volts	119
10-5	Edot vs. E phase plane, t_d = 0, k_t = .02, V = 1.5	119
10-6	C, Cdot, E vs. time of system with t_d = .5, k_t = .02, V = 1.5	120
10-7	Cdot vs. C plane of system with ramp input t_d = .5, k_t = .02, V = 1.5	121
10-8	Edot vs. E plane, ramp input, t_d = .5, k_t = .02 V = 1.5	122
10-9	C, Cdot, E vs. time, ramp input, t_d = .5, k_t = 2.0, V = 1.5	123
10-10	Cdot vs. C plane, t_d = .5, k_t = 2.0, V = 1.5	124

Figure	Title	Page
10-11	Edot vs. E plane, $t_d = .5$, $k_t = 2.0$, $V = 1.5$	124

TABLE

8-1	Data for first overshoot t_p and M_{p_t} under various time delays for a step input.	79
8-2	Data for steady state oscillation frequency vs. time delay	94
8-3	Steady state error vs. time delay	94
9-1	Optimum value of k_t for eliminating the effect of time delay	104
9-2	Calculated response of relay servo system Forcing function = unit step, $t_d = 0$	113
9-3	Calculated response of relay servo system Forcing function = unit step, $t_d = .5$ sec.	114

LIST OF SYMBOLS

A_{osc}	Oscillation amplitude
c, θ_o, e_o	Output signal
c, v	Rate output, or velocity
e_i, e, R	Input forcing function
E	Error
e^{-sT}	Transfer function of transportation lag
f	Frequency in cycle per second
G_D, N	Nonlinear element
G(s)	Open loop transfer function
H(s)	Transfer function of feedback element
K	Constant or system gain
k_t	Tachometer gain
M_{pt}	Magnitude of peak overshoot
T, t_d	Delay time
t_p	Time to peak overshoot
t_r	Rise time
ω	Frequency in rad/sec
ϕ	Phase angle

1. Introduction.

Systems which utilize feedback for control purpose have become essential elements in modern technology. They range all the way from simple toys to the most complex automatic factories and production equipment. Feedback control is the guidance that makes modern automation possible, and as such is in a large measure responsible for the ever-increasing productivity and rising standard of living of man.

It is noted that feedback control systems are seldom truly linear, but rather they are subject to the introduction of a considerable number of nonlinearities. Some nonlinearities may be introduced intentionally by the system designer. An example of this might be the selection of a relay servomechanism with its attendent nonlinearity. On the other hand there may be no way of avoiding a nonlinearity in the system design. An example of this would be the mandatory inclusion of a gear train with its attendent backlash. A nonlinearity which frequently can neither be avoided nor neglected is transportation lag.

In some control systems, especially in the field of process control, a non-eligible propagation time is encountered. That is, there is a definite lapse of time between the initiation of a change in the signal at one point in the system and the appearance at some other point of any effect of this change. But be sure that this kind of lag is not to be confused with the type of lag introduced by factors of the form $1/(TS + 1)$ or $1/(T^2S^2 + 2\zeta TS + 1)$ in the transfer function relating the two signals. Factors such as these in the transfer function of a component do reveal a tendency of the component to retard the signal, in a sense, by introducing phase lag at all frequencies, but they do not result in a finite dead time which must elapse before any effect of a change in input can be noted in the output of the component.

Dead time of appreciable magnitude can be observed, for example in long pipes carrying fluid under pressure. A change in pressure at one end of such a pipe is transmitted with finite velocity to the other end, possibly but not necessarily being distorted in transmission. Another type of system in which dead time is often encountered is shown in Fig. 1-1. In such a system two different fluids are mixed at the mixing point, and the relative concentration is read at the metering point. If the metering point is located at an appreciable distance from the mixing point, the dead time between the establishing of a change in relative concentration and the detection of the change may be appreciable.

Another example is illustrated below. In Fig. 1-2 the shaded area represents a slug of water at position 1, the outlet of the steam nozzle. Position 2 represents where the temperature sensing element will be placed. Let the temperature of the shaded area water slug at the steam nozzle be quickly changed from θ_1 to θ_2. Before the change can be measured, the slug of water must travel from the steam nozzle to position 2 where the sensing element would be located.

The effect of the delay caused by the slug traveling down the pipe can be shown by locating sensing elements at position 1 and 2 as shown in Fig. 1-3. If the time constants of the sensing element at position 1 is taken as the input and the element at position 2 as the output, Fig. 1-4 becomes the reaction curve for the process.

This type of reaction curve is typical of a dead time lag (or transportation lag) and may be defined by the time lag T. The magnitude of the dead time lag will depend on the water velocity and the distance between position 1 and 2. The lag time may be calculated as follows:

$$T(\text{min.}) = d(\text{ft}) \ / \ V(\text{ft/min})$$

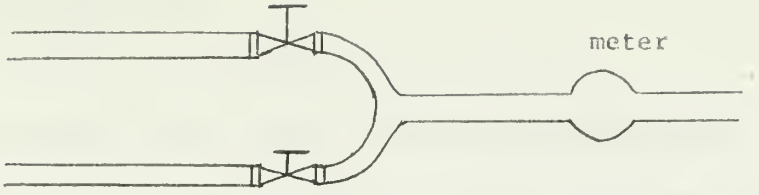


Fig. 1-1

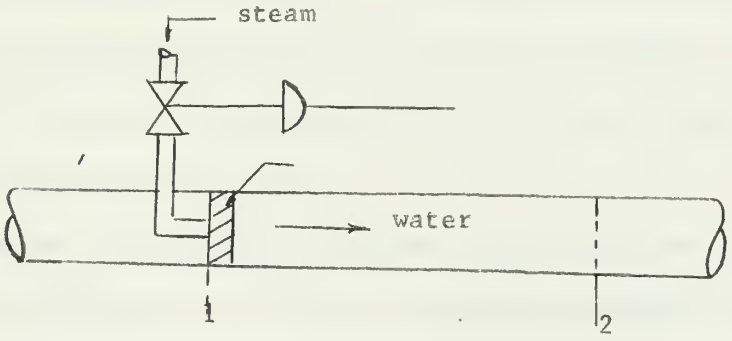


Fig. 1-2

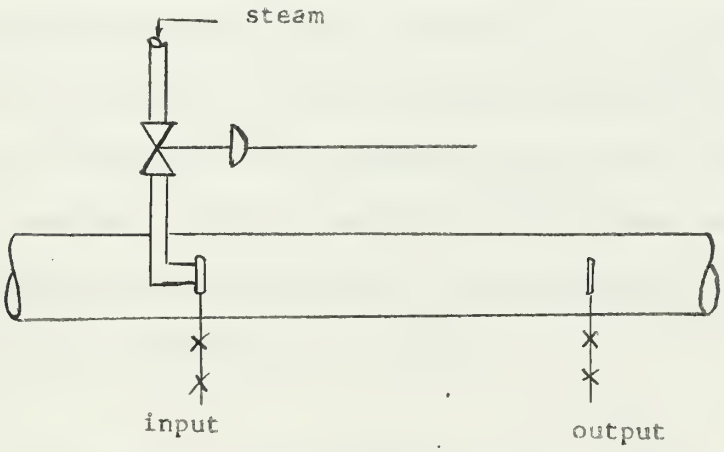


Fig. 1-3

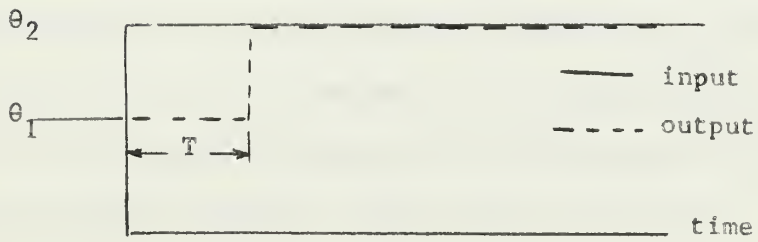


Fig. 1-4

where, T = dead time lag (min)

d = length (ft)

V = velocity of flowing medium (ft/min)

Dead time lag is very common in the process industries. Because of the complete lack of the response during the dead time, this type of lag is very detrimental to control. An examination of the equation defining the dead time lag shows that increasing the length or decreasing the velocity will increase the lag time, and we shall see later, make the problem of control more difficult.

To determine the frequency response of the dead time lag, a sinuscilal wave having an amplitude of 10°F will be generated around a bias temperature θ_1. At some low frequency the curve as recorded by the sensing elements at position 1 and 2 would appear as shown in Fig. 1-5(a).

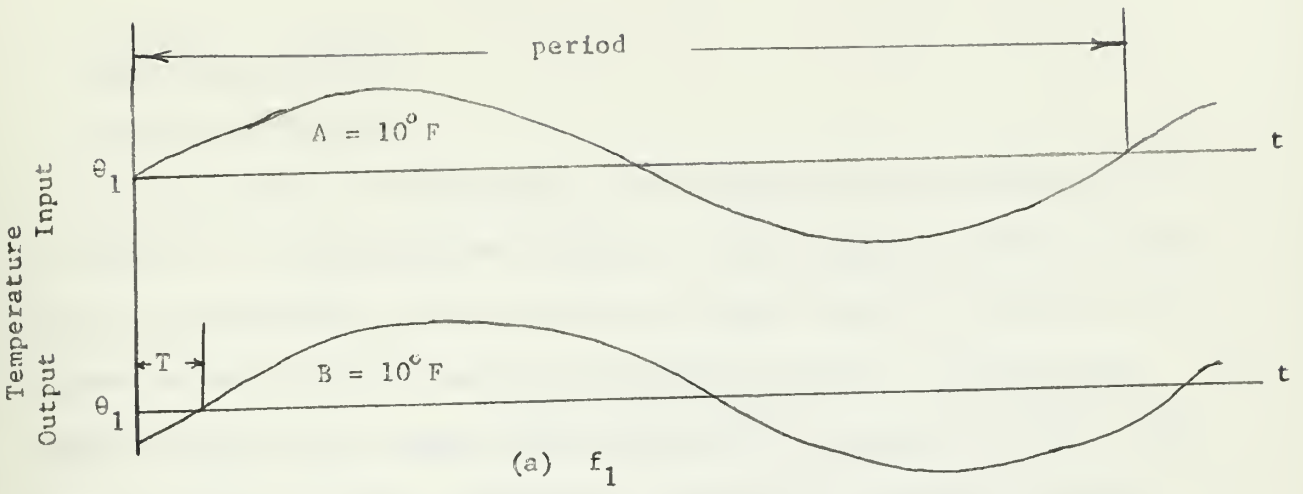
Fig. 1-5(b) and 5(c) show similar results for two successively higher frequencies. As can be seen, as the frequency increases the period decreases. However, the lag time which is dependent only on the velocity and length remains constant. Thus the magnitude of the phase angle will increase proportionally to each decrease in period (increase in frequency). The phase angle may be calculated as follows:

$$\phi \text{ (degree)} = \frac{-T(\text{min})}{\text{Period (min/cycle)}} \times 360 \text{ (degree/cycle)}$$

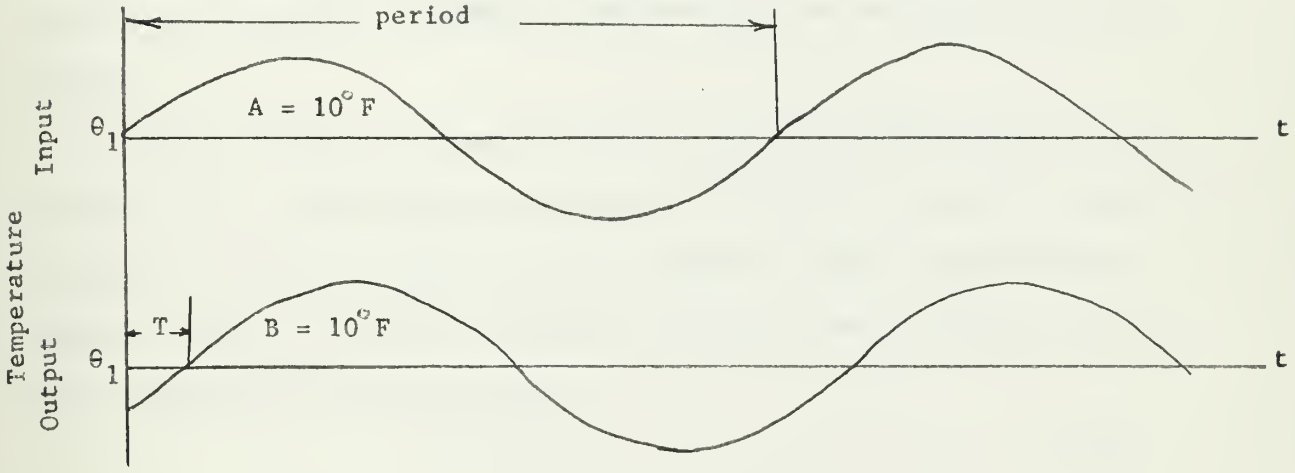
or

$$\phi \text{ (degree)} = - T(\text{min}) \times f(\text{cpm}) \times 360 \text{ (degree/cycle)}$$

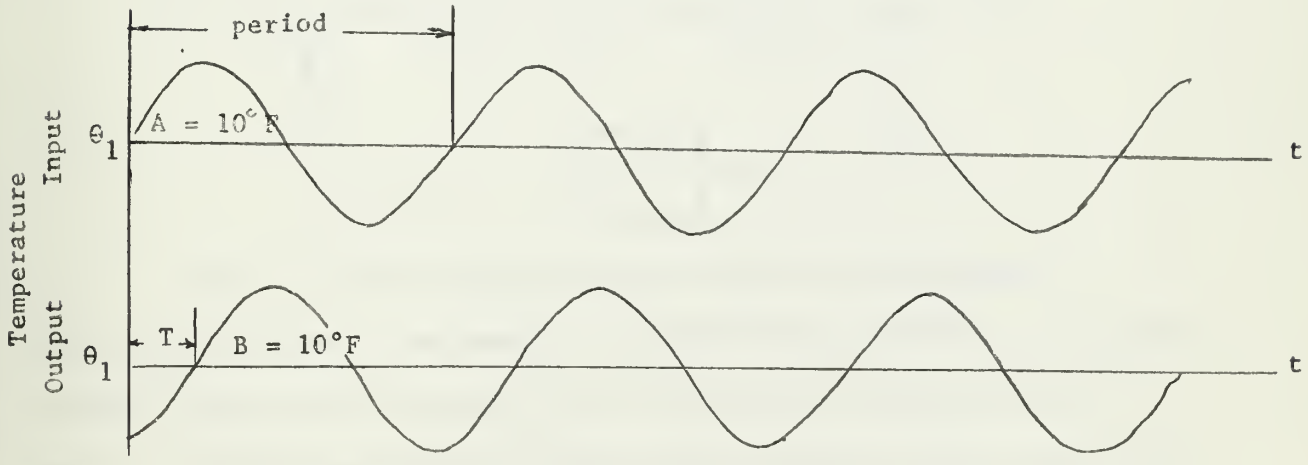
From this equation we see that the phase angle for the dead time lag is directly proportional to the frequency. The phase angle approaches zero degrees as the frequency approaches zero and negative infinity as the frequency approaches infinity. This will be discussed in § 6.



(a) f_1



(b) $f_2 > f_1$



(c) $f_3 > f_2$

Fig. 1-5

2. Relay servo system.

2.1 General Description.

When economy is of importance, many feedback control systems utilize a contactor for energizing the actuator. Such a system is called a contactor or, alternatively, an on-off control system, since the contactor is an element whose output can usually have only two values. In some cases these correspond to a fixed positive action and no actuator action. This usually is the case with thermostatic controls, for example. In other cases the two values correspond to maximum positive and negative actuator action.

In Fig. 2-1 it is assumed that N is a relay or switch which has an output +1 or -1, depending on the sign of E = R - b. A positive output corresponds to a positive sign of E and a negative output corresponds to a negative sign of E. When the switch is a perfect switch this can be denoted symbolically by the equation

$$V = \text{Sign } E \tag{2.1}$$

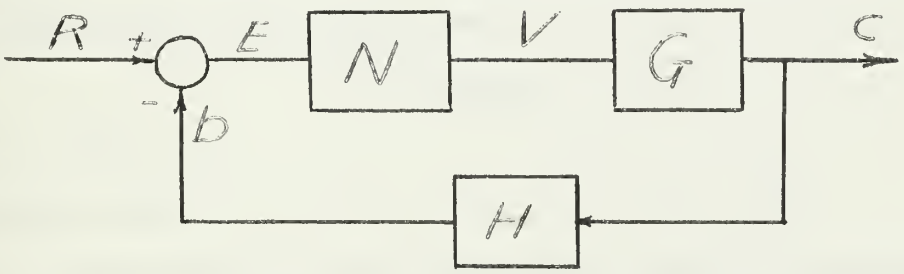


Fig. 2-1 Block diagram of the contactor servomechanism

Since the contactor is frequently an electrically controlled or equivalent element, there may be a delay in the initiation of switching action as well as a finite time interval during which switching action takes place. The contactor action ordinarily applies full control force or torque to the system. Since the output of N is therefore not necessarily unity, the

actual value is included in the gain constant of the following element.

The block G in Fig. 2-1 will be assumed to be a linear servomotor with viscous damping and no load torque except that due to inertia. The limiting case of zero viscous damping is of interest and will also be considered. Including the gain constant of the switch, the motor transfer function is

$$G = \frac{T_o}{JS(S + 1/\tau_m)} \tag{2.2}$$

where $T_o = KE_o$ is the stall torque of the motor and J is the total inertia of motor and load. In the absence of damping, $1/\tau_m$ is set to zero in this equation. It should be emphasized that since N is nonlinear element, there is no particular reason to assume that the feedback element is linear. In particular, the feedback signal could be any algebraic function of position and its derivatives. However, for the method of analysis used here the feedback function can depend only on the position and velocity and must not on the derivative of the velocity (acceleration). The symbol H, therefore, no longer represents a linear transfer function, but a possibility nonlinear feedback function involving the output position and velocity. Symbolically this will be written as

$$b = f(C,\dot{C}) \tag{2.3}$$

where $C = dC/dt$

2.2 Relay servos

The system behavior can be prescribed by combining the equations:

$$\begin{aligned} \ddot{c} + \frac{1}{\tau_m}\dot{c} &= +\frac{T_o}{J} \qquad R - f(c,\dot{c}) > 0 \\ \ddot{c} + \frac{1}{\tau_m}\dot{c} &= -\frac{T_o}{J} \qquad R - f(c,\dot{c}) < 0 \end{aligned} \tag{2.4}$$

Between switching times the response is described by a second order differential equation. For this reason such a system is called a second-

order contactor servo or feedback system. The two equations are integrable between switching times and the entire solution can be obtained by matching final and initial conditions of the two sets of solutions at the proper switching time. The solution is therefore easy to obtain in principle. The details, however, are so complex that another method of solution is desirable.

The analysis of the system can be simplified by using the fact that it is possible to plot the response of the system as a trajectory in the phase plane, the trajectory being composed of a series of arcs or curves which have a characteristic shape dependent upon the properties of the system. The behavior of a physical system where only one coordinate is of interest is usually specified by determining this coordinate as a function of time, $C = f_1(t)$. This function depends upon the differential equation describing the system and upon a set of initial conditions. For some purpose, an equal suitable way of describing the behavior of the system is by means of a functional relationship between position and velocity:

$$\dot{C} = f_2(C) \tag{2.5}$$

When such a relation is known, the time dependence of C, if required, can be found directly by integrating the equation

$$\frac{dc}{dt} = \dot{c} = f_2(c) \quad or \quad t - t_0 = \int_{c_0}^{c} \frac{dx}{f_2(c)} \tag{2.6}$$

Equation 2.5 can be represented graphically with position and velocity taken as rectangular cartesian coordinates; position is usually plotted as abscissa and velocity as ordinate. The behavior of the system is described by a plot of all successive pairs of values of position and velocity. The plane on which this plot is made is called the phase plane, and

the graph of Equation 2.5 is called the phase-plane trajectory.

For the system to be considered in this paper, the differential equations of interest will not contain time explicitly. For the second-order system this implies that the acceleration can be written as

$$\ddot{C} = F(C, \dot{C})$$

The importance of this concept here results from the fact that for such systems, the behavior and therefore the trajectory are uniquely specified by the initial values of C and $\dot{C}$. This means that there is an only trajectory through a given (nonsingular) point in the phase plane. In this paper it is restricted to the systems described by second order differential equations and hence by phase plane trajectory in the two dimensional C-$\dot{C}$ phase-plane.

The slope of the trajectory in the phase plane can be written as

$$\frac{d\dot{C}}{dC} = \frac{\frac{d\dot{C}}{dt}}{\frac{dC}{dt}} \qquad (2.7)$$

According to this equation the slope is equal to the quotient of acceleration and velocity. In the upper half plane, $\dot{C}$ is positive and nonzero, and therefore the slope is nonifinite and either positive negative, depending upon whether the acceleration is positive or negative. Similarly, in the lower half plane the slope is non-infinite and either positive or negative, depending on whether the acceleration is negative or positive. For a point on the horizontal axis the velocity is zero, and therefore unless the accelerations also zero the slope $d\dot{C}/dC$ is infinite. In general, then, trajectories cross the horizontal axis vertically. Direction of motion along the trajectory can always be determined by the fact that for $\dot{C}$ positive C must become more positive with time and

for $\ddot{C}$ negative C must become more negative with time. Motion along the trajectory is therefore to the right in the upper-half plane and to the left in the lower-half plane. The horizontal component of the trajectory velocity is equal to the system velocity $\dot{C} = dC/dt$ and the vertical component of the trajectory velocity is equal to the system acceleration $\ddot{C} = dC/dt$. These considerations make it clear that a trajectory can not end except on the horizontal axis. If there are points on the horizontal axis where both $\dot{C}$ and $\ddot{C}$ are zero they are called singular points, and the value of C at such a point corresponds to an equilibrium condition of the system. Trajectories can emerge from or return to some of these singular points with a non-infinite slope, since the slope $d\dot{C}/dC$ is undefined at this point except through other considerations.

2.3 Phase plane trajectory.

Some of the ideas in §2.2 will now be illustrated with the system shown in Fig. 2.1 and defined by Equation 2.2 through 2.4. The Equation 2.4, which are second-order differential equations, can be reduced to pairs if first-order equations by substitution of Equation 2.6

$$V = \dot{C} = \frac{dC}{dt}$$

which gives

$$\frac{dV}{dt} + \frac{1}{\tau_m} V = \pm \frac{T_o}{J} \tag{2.8}$$

an alternative expression for velocity, $dV/dt = V(dV/dC)$, can be used with Equation 2.8 to give the following equation:

$$V \frac{dV}{dC} = - \frac{V}{\tau_m} \pm \frac{T_o}{J} \tag{2.9}$$

This equation can now be integrated with the result

$$C - C_o = -\tau_m \left[V \pm \frac{J_m T_o}{J} \log\left(1 + \frac{JV}{\tau_m T_o}\right) \right] \tag{2.10}$$

Evidently the constant of the integration C_o is the value of C at which V = 0. This equation can be put in dimensionless form by the substitutions

$$\bar{v} = \frac{J}{\tau_m T_o} V , \qquad \bar{C} = \frac{J}{\tau_m^2 T_o} C \tag{2.11}$$

which gives

$$\tau - \tau_o = -\left[\bar{v} \pm \log (1 \mp \bar{v})\right] \tag{2.12}$$

Typical trajectories for particular values of $\bar{C}_o$ are shown in Fig. 2.2. Through use of the dimensionless variables $\bar{v}$ and $\bar{c}$, all systems of this type have identical trajectories.

In the absence of damping, the trajectories are parabolas. If $1/\tau_m$ is set to zero, integration of Equation 2.8 gives the result

$$C - C_o = \pm \frac{J}{T_o} \frac{V^2}{2} \tag{2.13}$$

In this case the substitution 2.11 can not be used. For the case of finite damping it is sometimes convenient to have an expression for the trajectory which shows the deviation from the parabolic shape more directly than Equation 2.10 or 2.12. This can be obtained by use of the power series expansion for log (1 + x):

$$C - C_o = \pm \frac{J}{T_o} \frac{V^2}{2} \left[1 \pm \frac{2}{3} \frac{JV}{\tau_m T_o} + \frac{2}{4} \left(\frac{JV}{\tau_m T_o}\right)^2 \pm \cdots\right] \tag{2.14}$$

or

$$\bar{C} - \bar{C}_o = \pm \frac{\bar{V}^2}{2} \left[1 \pm \frac{2}{3}\bar{V} + \frac{2}{4}(\bar{V})^2 \pm \frac{2}{5}(\bar{V})^3 + \cdots\right] \tag{2.15}$$

These equations are particularly useful when an approximate result is desired. For accurate computations Equation 2.10 or 2.12 is more useful, since the series given above converges rather slowly except for small values of $\bar{v}$, and indeed diverges for $|\dot{v}| > 1$.

Reference to Fig. 2.1 and Equations 2.1 and 2.3 shows that the

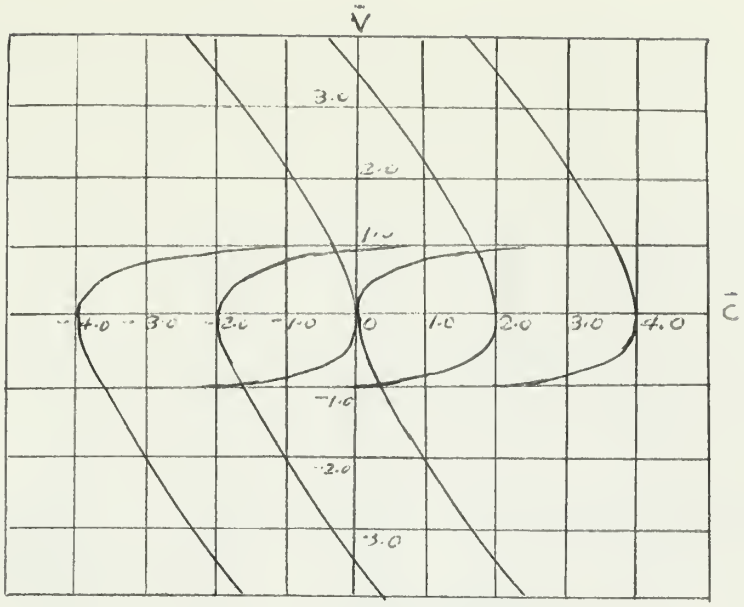


Fig. 2.2. Phase plane trajectory for the second-order contactor servomechanism.

condition governing switching from one trajectory family to the other is given by the equation

$$R - f(c, \dot{c}) = 0 \tag{2.16}$$

For the present it will be convenient to restrict attention to step inputs with initial velocity zero. In this case a simple translation of the origin can be used so that the initial conditions are $c = -c_1$, $V = 0$, and the final desired steady-state conditions are $R = C = V = 0$. Setting $R = 0$ in Equation 2.16 gives the equation

$$f(c, \dot{c}) = 0 \tag{2.17}$$

The graphical plot of this equation separates the phase plane into two parts, in one of which the system follows trajectories corresponding to $V = +1$, and in the other of which the system follow trajectories corresponding to $V = -1$, this is illustrated in Fig. 2.3.

For certain functions $f(c,\dot{c})$ an anomalous situation arises. Consider, for example, the point P_1 in Fig. 2.3, which will supposed to be reached by motion along the trajectory marked (1). When the switching line is

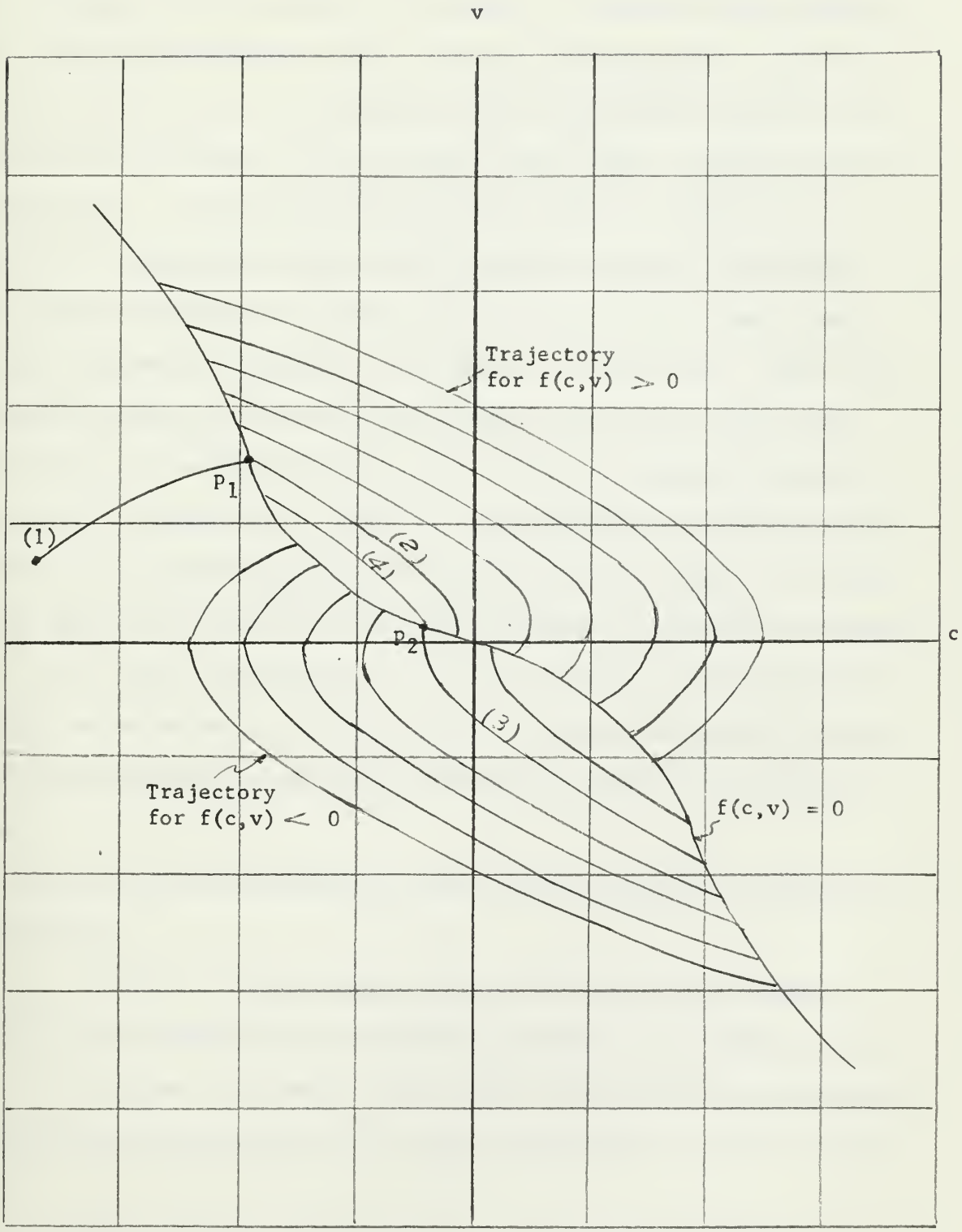


Fig. 2-3. Switching line for contactor servomechanism.

reached, the system will normally transfer to the trajectory of the other family which pass through P_1, this is the trajectory (2) in Fig. 2.3. However, when the point P_2 is reached, for example along trajectory (3), a paradoxal situation arises. Although switching should take place at this time, there is no trajectory of the other family along which the system can move, since P_2 is an end point rather than a starting point for trajectory (4). In terms of real physical systems the paradox can be resolved by recognizing that a small delaying switching is unavoidable. This permits the system to overshoot the switching line so that a small section of allowable trajectory is available in other family. When this trajectory reaches the switching line another slight delay allows overshoot back into the left-half plane. The resultant trajectory consists of a series of short arcs in which the system oscillates about the switching line as it approaches the equilibrium point at the origin. In terms of the ideal system, one can say that the trajectory follows the switching line to the origin accomplished by an infinite number of switching which provide, on the average, the proper force to give an acceleration consistent with the location in the phase plane and the slope of the switching line.

Using the above considerations some interesting general conclusions can be drawn:

(1) For no viscous damping ($\tau_m = \infty$) and no rate feedback [$f(c,\dot{c}) = c$] the trajectories are parabolas separated by the vertical line c = 0. The system is neutrally stable in the sense that for any amplitude of step input the oscillation will persist indefinitely, as shown in Fig. 2.4a.

(2) For viscous damping and no rate feedback [$f(c,\dot{c}) = c$], the

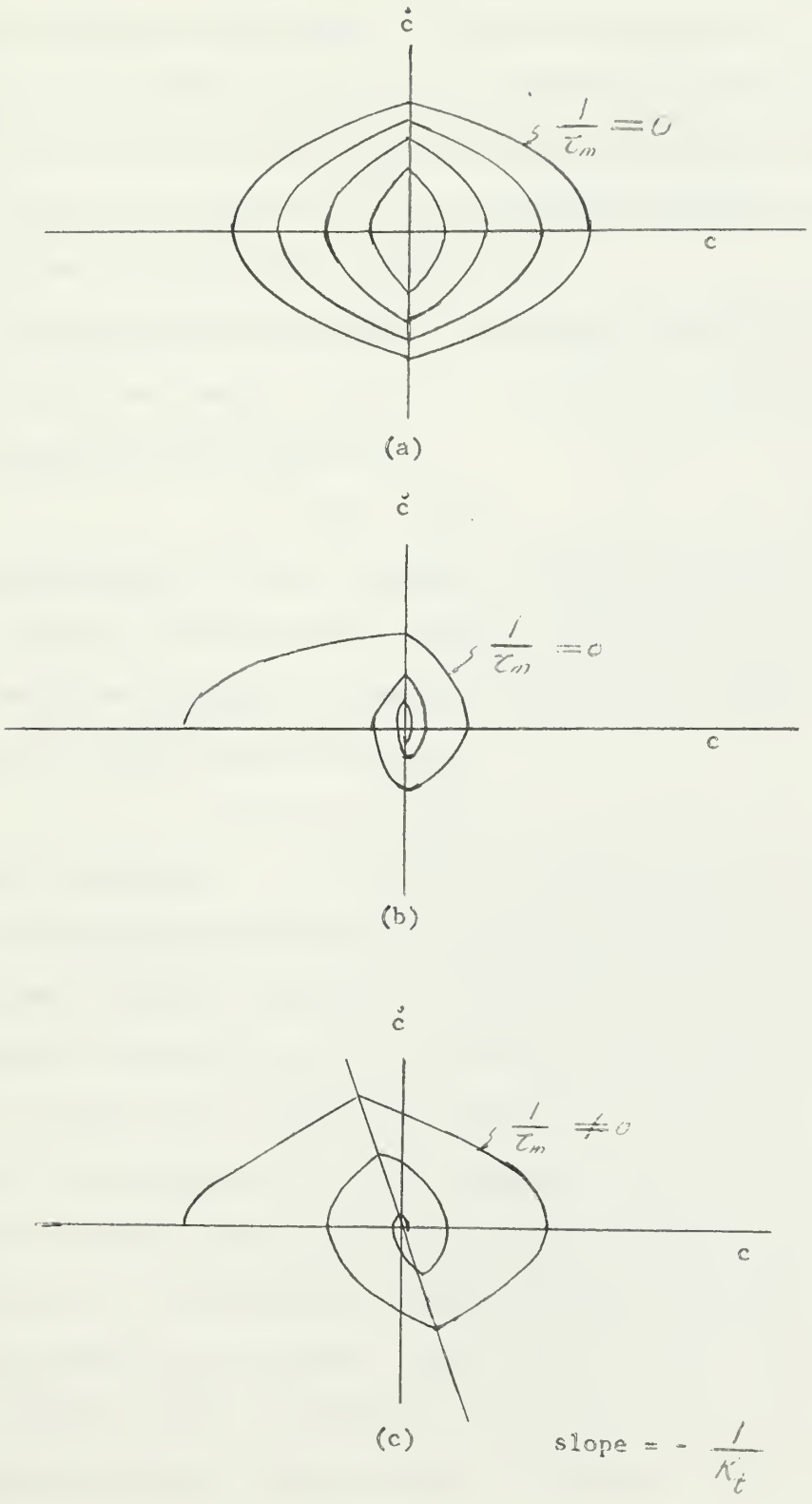


Fig. 2-4. Typical phase plane trajectories

trajectory are the "asymmetric parabolas" of Fig. 2.2, separated by the verticle line c = 0. The system exhibits a damping or decrement of the oscillations produced by a step input. The per-unit decrement is a function of amplitude and approaches zero as the amplitude approaches zero. In other words, for large input the initial per-unit decrement is high, but decreases to zero as the amplitude of the oscillation approaches zero. This can be observed qualitatively in Fig. 2.4b.

(3) With linear rate feedback, the feedback function is $f(c,\dot{c}) = c + K_t V$ If there is no viscous damping ($\tau_m = \infty$) , then the trajectories are parabolas separated by the straight line

$$\dot{c} = -\frac{1}{K_t} c \qquad (2.18)$$

Again the system shows a per-unit decrement or damping which varies with amplitude. However, in this case the damping is small for large amplitude and as the amplitude decreases the damping increases until the final trajectory into the origin lies along the line

$$\dot{c} = -\frac{1}{K_t} c$$

with no further overshoots. An illustration of this is shown in 2.4c.

The above conclusions shows that no simple linear feedback function will give optimum response, regardless of the system's viscous damping. If rate feedback is adjusted to give no overshoot for the largest expected input then the system moves to the origin slowly for small inputs. If the rate feedback is adjusted to give faster response for small input, then overshoot and oscillations will follow application of a large step input.

It is natural to ask whether any feedback function can be found to give optimum response for all inputs. This turns out to be a difficult question. Since the system is nonlinear and does not have an easily determined response for arbitary inputs. However, a definite answer can be found

for the case of step inputs by following chain of reasoning. For most rapid response without overshoot it is necessary to apply full positive acceleration torque until that instant at which full negative torque will bring the system to rest at the origin in the phase plane. This is illustrated in Fig. 2.5 for a particular initial starting point. The trajectory shown passing through the origin can be designated as the homing trajectory. In order to achieve switching on this trajectory for all inputs, it is necessary that feedback function set equal to zero,

$$f(c, \dot{c}) = 0 \tag{2.19}$$

being the equation of this homing trajectory. In this case, when any trajectory crosses this line switching will occur, so that the application of full reverse torque will just bring the system to reset at the origin.

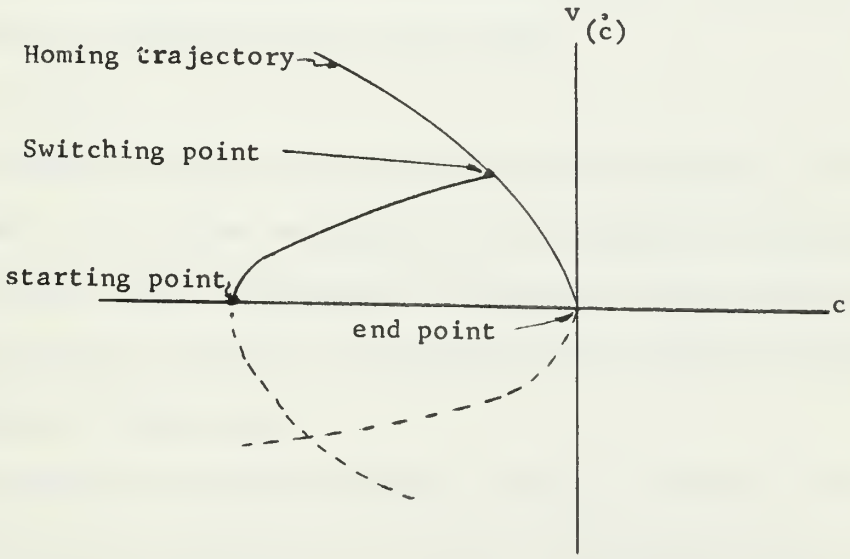


Fig. 2.5 Switching point for optimum step response

Since the trajectory of Equation 2.8 is roughly parabolic, a suitable form of the feedback function is

$$f(c, v) \approx K(\bar{c} + \tfrac{1}{2}\bar{v}|\bar{v}|) = c + \frac{J}{2T_o} v|v| \tag{2.20}$$

Strictly speaking, this equation should be modified to match the appropriate "asmmetric parabola", although for practical purposes the refinement is not necessary. In other words, if the position feedback is linear and the rate feedback roughly quadratic (except for sign) then it is possible to obtain optimum response for this particular type of input. It will be seen that switching delays give rise to oscillatory motion of small amplitude unless a linear velocity-feedback term is introduced. For this reason it is usually desirable to introduce a small amount of such feedback (in addition to the quadratic feedback) in order to eliminate or minimize this oscillation. This does not significantly affect the results for large amplitudes.

It is quite important to remember that this optimization does not necessarily give optimum response for arbitary inputs. Further consideration is beyond the intended scope of this paper.

2.4 On-off control.

There probably are more on-off discontinuous controllers in the world than the number of all other servomechanisms and regulators put together. Hot-water heaters, ovens, refrigerators, furnaces, electric irons, battery chargers, liquid-level controls, and compressed-air pressure regulators are all commonly of this type.

A relay system is one which contains a decision function. It is commonly in the error amplifier and in its most elementary form is an on-off switch. If the positive error exceeds a certain amount, the switch is turned from on to off, and if the negative error exceeds a certain amount, the swtich is turned from off to on. When the switch is on, a predetermined constant corrective force is applied into the rest of the system to bring the output back into correspondence with the desired value. When the switch

is off, this force is removed. The differential equation for each is the open-loop equation, with one term, the corrective force, having two different values for the two periods. Each switch operation starts a new period for which the initial values are the final values of the previous period.

One characteristic of a relay is that it delivers an output having one of only two values, independent of the continuous changes in the input, except they specify the switching instants. It also has hysteresis, the phenomenon that the input values for switching on and for switching off are not exactly equal. The actual time of motion of the armature will be considered negligible, but the system may contain the dead time of flow of material through a long pipe, equivalent to the time of transmission of an electric signal on a long distoritonless line.

The term relay is used to apply to any device having at least the first of these characteristics. In this paper, we are interested primarily in discontinuous control of continuous processes, so that the relay is only one part of the system.

The on-off control, or two position control, is undoubtedly the most widely used type of control for both industrial and domestic service. It is the kind of control generally employed on home heating systems and domestic water heaters.

Two position control is a position type of controller action in which the manupulated variable is quickly changed to either a maximum minimum value depending upon whether the controlled variable is greater or less than the set point. The minimum value of the manupulated variable is usually zero (off). This mode of control is illustrated by the electric level-control in Fig. 2.6. A float in the vessel operates an electric switch which controls power to a solenoid valve. When the liquid level rises,

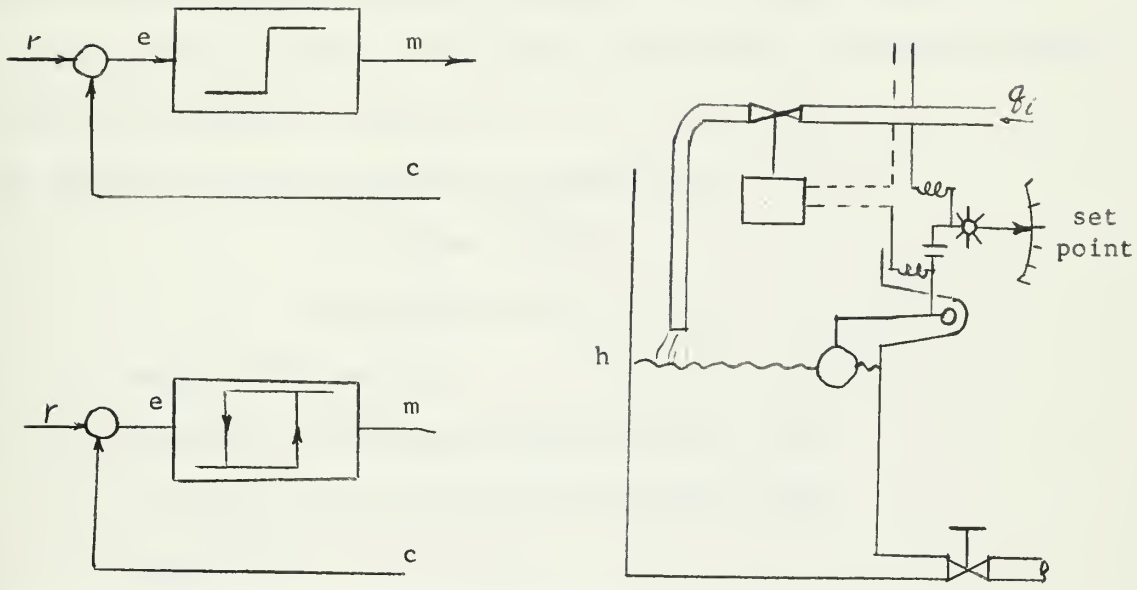


Fig. 2-6. Two position control

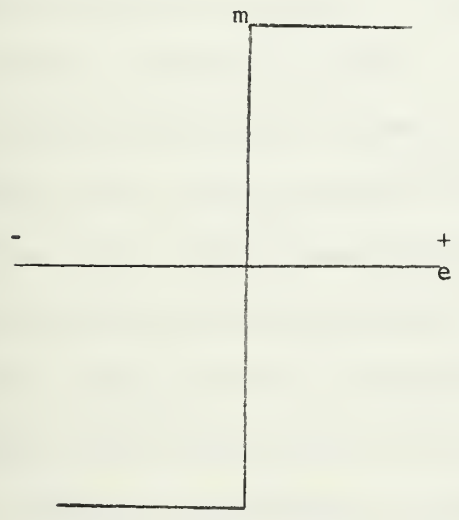


Fig. 2-7
Two-position control without differential gap

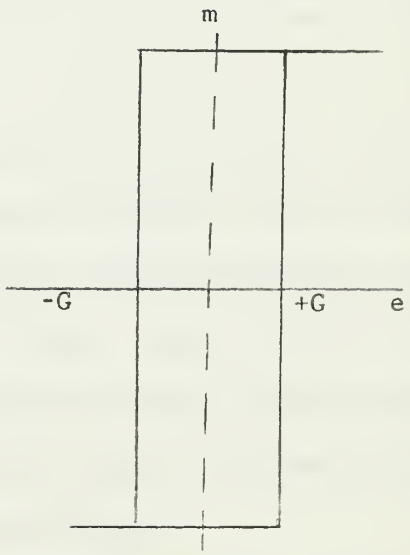


Fig. 2-8
Two-position control with differential gap

the switch contacts are closed, the solenoid valve closes, and the flow is cut off. When the liquid level falls, the switch contacts are opened, the solenoid valve opens, and the inflow resumes. If the float lever has no bearing friction and the electrical contact draw no arc, the action is sharp or "knife'edge" as shown in Fig. 2-7.

The equation for the two-position control are

$$m = M_1 \text{ when } e > 0$$

$$m = M_o \text{ when } e < 0$$

where m = manupulated variable

M_1 = maximum value of manupulated variable (on)

M_o = minimum value of manupulated variable (off)

e = deviation

Thus two-position control must be described by two equations, each applying in a certain region of deviation.

A differential gap in two-position control causes the manupulated variable to maintain its previous value until the controlled variable has moved slightly beyond the set point. In actual cperation it is the same as hysteresis, as may be seen in Fig. 2-8.

A differential gap is caused in the two-position controller of Fig. 2.6. If small static friction exists at the bearing on the float arm, the liquid level must rise slightly above the desire value to create sufficient buoyant force to overcome friction when the level is rising. Also, the liquid level must fall slightly below the desired value when the level is falling so that the weight force may overcome the friction. This kind of differential gap may be caused by unintentional friction and lost motion.

A differential gap may be intentional, as when a magnet is installed on the float arm in Fig. 2.6, causing a hysteresis in float-arm action. Similar arrangements are common in domestic thermostats and are employed for the purpose of preventing rapid operation of switches and solenoid valves for reducing of electrical contacts.

3. Description of the system.

The nonlinear feedback control system to be studied is a relay servo system as shown in Fig. 3-1.

Block one contains the transportation lag. The nature of this block and its analytic representation will be discussed in section 6.

Block two is the ideal relay, no dead zone and no hysteresis.

Block three is the motor load unit or plant. It is assumed to be a second order system. In order to have the transfer function always low pass in nature, we choose the value of G(s) as

$$G(s) = \frac{1}{s(s+1)} \tag{3.1}$$

Block four is a tachometer. In this system velocity feedback compensation is used.

This system, in addition to the velocity feedback, will also use unity position feedback.

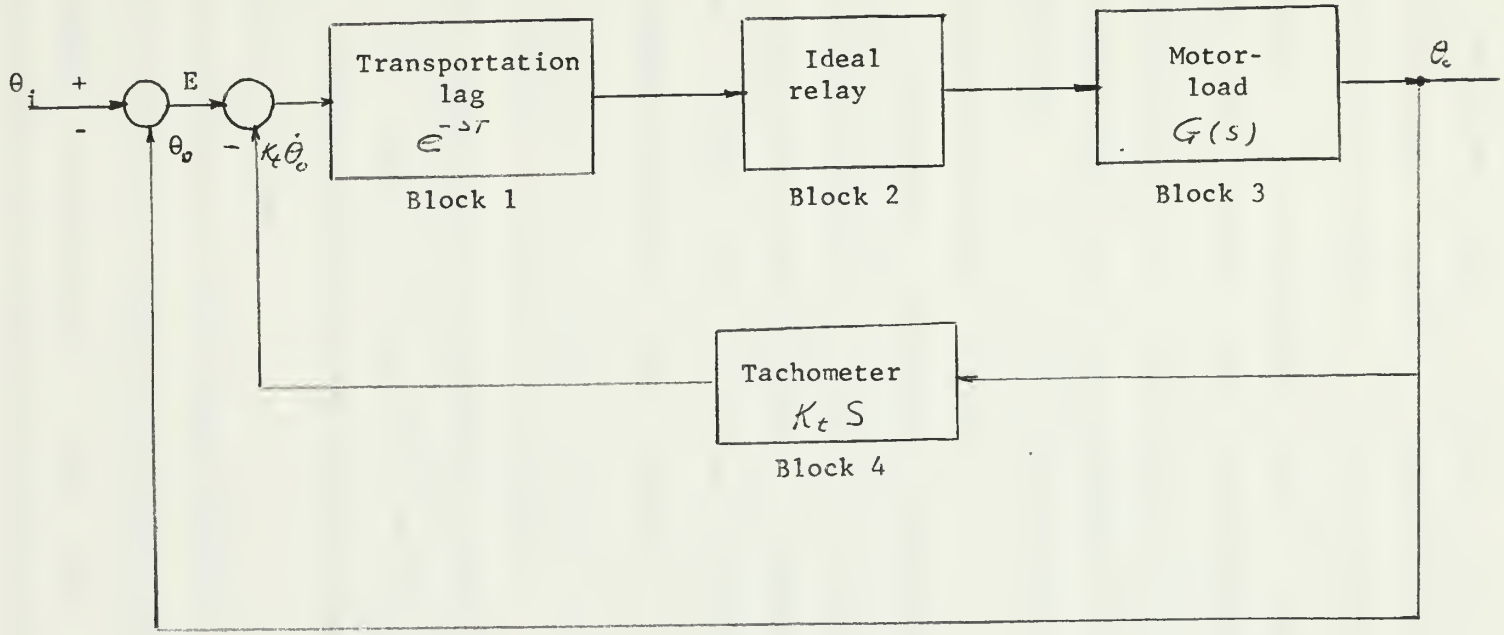


Fig. 3-1 Generalized block diagram of basic system

4. Nonlinear analysis.

4.1 General

In many cases the system designer is forced to work with components which are not entirely linear. In order to linearize the response of a nonlinear actuator we use the feedback. The basic difficulty with the nonlinear equation is that it is not possible to exhibit a general solution. Associated with this is the fact that there is no complementary function or transient solution made up of terms with characteristic shapes, there are, therefore, no system roots such as were defined for linear systems with constant coefficients. It is not the scope of this paper to include a comprehensive treatment of nonlinear systems.

Many systems are approximately linear when the amplitude of the system variables are not too large or too small. For large enough amplitude all physical systems must become nonlinear. A system which for small amplitude is a linear unstable system must become nonlinear when the oscillatory amplitude exceeds some value. Such a system will sometimes reach a stable oscillatory motion which is termed a limit cycle. Although such motion is usually periodic, it need not be sinusoidal.

There are other large-amplitude nonlinearities which are in a sense destablizing. In these systems, small amplitude response may be stable and satisfactory, but large disturbances may send the system into uncontrolled large-amplitude oscillations.

There are also nonlinearities which become important for small amplitude motions. Mechanical friction and backlash are good examples of this type. In the case of backlash it is not unusual to find a limit cycle of small amplitude which leaves the system capable of following commands or inputs

with good accuracy and in an essentially linear mannes except for the small limit cycle motion which is superimposed on the desired response.

4.2 The describing function.

Many functions involved in automatic control are not desirable by elementary techniques, and they must be described by graphical or other methods.

The output of a nonlinear element will not necessarily be sinusoidal even though the input is a pure sinusoid, but it will generally contain fundamental and harmonic waves. The frequency of the fundamental of the output will often be the same as the frequency of the input. In this case a describing function can be used and the element may be treated as a linear element.

Since most elements of a feedback system are of a low-pass nature, the higher harmonics may be greatly attenuated in the rest of the loop, so that the input to the nonlinear element may be quite accurately sinusoidal. If this is the case, it is necessary to compute the transmission around the loop for the harmonics. In particular, for the nonlinear element we need only a function to describe the fundamental component in the output as a function of the amplitude and frequency of the sinusoidal input. This function has been named a describing function. Since the describing functions are approximations to the output-input ratio of an element having slight nonlinearily, it is anolagous to the transfer function of a linear system, and for certain purposes will be used in a similar manner. There is, of course, one important difference. Since a transfer function describes a linear system, it may be frequency-dependent but it is not amplitude-dependent. On the other hand, the describing function is usually amplitude-dependent and may also be frequency-dependent. The describing function is expressed in

terms of the amplitude and phase of the fundamental of the output signal related to the amplitude and phase of the input signal. The describing function can be obtained analytically or experimentally.

Describing functions are generally useful in automatic control analysis if the neglect of the harmonics in the output signal does not cause difficulties. They may be used with adequate results when the element in question is followed by one or more elements having appreciable attenuation of the harmonics in question.

4.3 Computation of the describing function.

Consider a nonlinear element whose input is sinusoidal,

$$x = X \sin \omega t \tag{4.1}$$

and whose steady state output can be written as a Fourier series:

$$\begin{aligned} Y = y(t) &= a_o/2 + a_1 \cos \omega t + b_1 \sin \omega t + a_2 \cos 2\omega t + b_2 \sin 2\omega t \\ &\quad + \cdots\cdots \\ &= a_o/2 - \sum_{k=1}^{\infty} (a_k \cos k\omega t + b_k \sin k\omega t) + \cdots \end{aligned} \tag{4.2}$$

Evidently

$$a_k = 2/\pi \int_o^{\pi} y(t) \cos(k\omega t) d(\omega t) \tag{4.3}$$

$$b = 2/\pi \int_o^{\pi} y(t) \sin(k\omega t) d(\omega t) \tag{4.4}$$

The describing function is defined as the complex number

$$N = N_1 + jN_2 = N_o \epsilon^{j\psi} \tag{4.5}$$

where

$$N_1 = b_1/X \qquad N_2 = a_1/X$$

or, alternatively,

$$N(j\omega) = G_D(j\omega) = \frac{\sqrt{a_1^2 + b_1^2} \; \angle \tan^{-1} \frac{a_1}{b_1}}{X} \tag{4.6}$$

This definition immediately puts certain restrictions on the permissible nonlinearities. The output function must be a periodic function with period equal to that of the input. According to this definition, devices which produce subharmonic response do not have a describing function.

The form of Equation 4.1 implies that the input is sinusoidal and does not contain d-c bias. If a nonlinearity in a feedback system is not symmetric, then in general there will be such a bias.

Because we are only interested in the nonlinear device of ideal relay only ideal relay will be discussed. The characteristic of the ideal relay is as shown in Fig. 4-1.

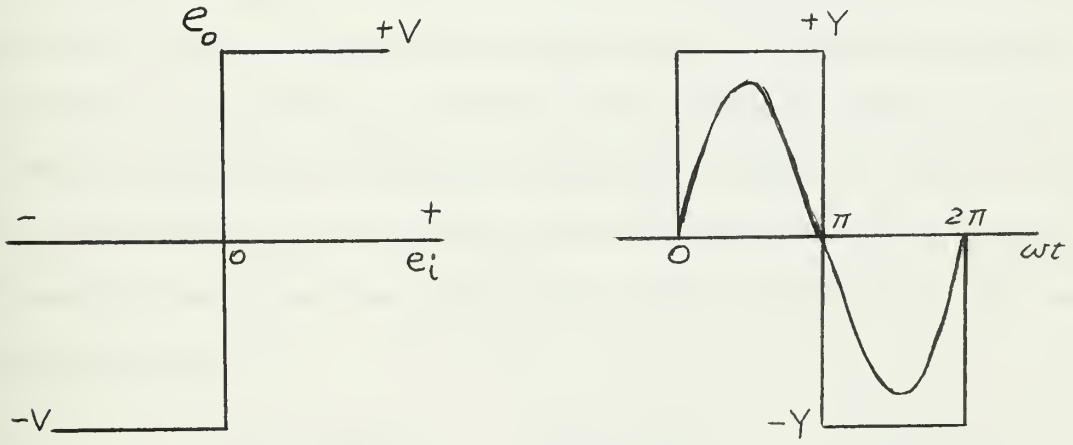


Fig. 4-1. Characteristic of ideal relay.

From Fig. 4-1

$$e_o = + Y \qquad o < \omega t < \pi$$

$$e_o = - Y \qquad \pi < \omega t < 2\pi$$

The describing function of the ideal relay can be obtained very simply by carrying the integration indicated in Equations 4.3 and 4.4 and 4.6. And in calculating the describing function, we are only interested in the fundamental terms. Thus we have the describing function of the ideal relay as

$$G_D(j\omega) = \frac{4Y}{\pi X} \qquad (4.7)$$

From the above results, we know that the describing function of the ideal relay is amplitude-dependent and independent of frequency. This function is plotted against X in Fig. 4.2.

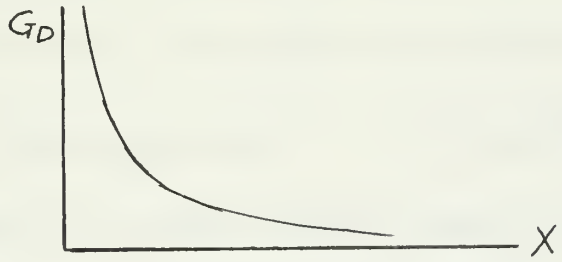


Fig. 4.2. Describing function of ideal relay.

4.4 Analysis of the limit cycle.

4.4.1 Frequency response method

Limit cycle angular frequency ω_L, as predicted by the frequency response method, is such that the loop phase shift is 360 degrees and this occurs when the loop frequency response function has no imaginary part.

The flow diagram for sinusoidal quantities is shown in Fig. 4.3; the relay is represented by a describing function in cascade with a time delay transfer function and the relay output takes on the values + or - unity. The condition for

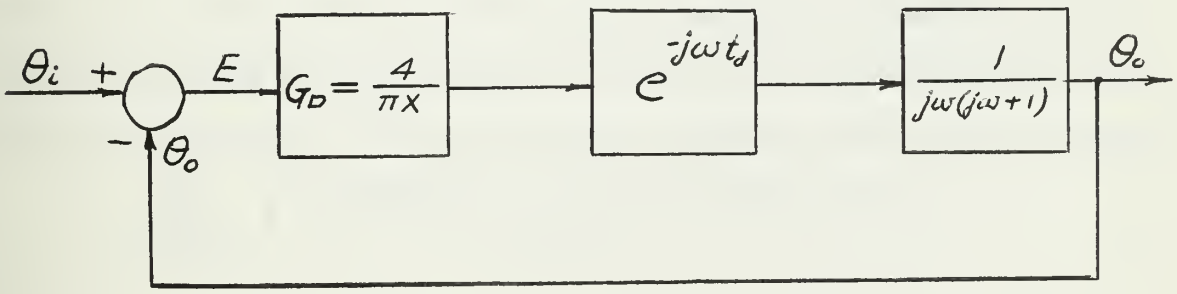


Fig. 4.3.

a limit cycle is therefore met when

$$G(j\omega) = \frac{e^{-j\omega_L t_d}}{j\omega_L(j\omega_L + 1)} \quad \text{is real}$$

by rationalizing this expression the condition becomes:

$$e^{-j\omega_L t_d}[\omega_L^2 + j\omega_L] \quad \text{is real}$$

Expanding the exponential term gives the limit cycle condition as

$$(\cos \omega_L t_d - j \sin \omega_L t_d)(\omega_L^2 + j\omega_L) \quad \text{is real}$$

Equating the imaginary terms to zero yields the condition

$$\omega_L \tan \omega_L t_d = 1 \qquad P = \frac{2\pi}{\omega_L} \tag{4.8}$$

For a given value of t_d (delay time) the dimensionless period $P = \frac{2\pi}{\omega_L}$ is multivalued. These values correspond to the possible modes of oscillation and only the longest period is stable.

According to the frequency response method the servo loop gain at limit cycle frequency is unity since the oscillation is exactly regenerated in the loop. To determine the limit cycle amplitude at the input to the on-off element it is therefore necessary to solve the equation obtained by equating the loop frequency response function to unity at the angular frequency, that is

$$\frac{4}{\pi X_L} \left\{ \frac{e^{-j\omega_L t_d}}{j\omega_L - \omega_L^2} \right\} = 1 \tag{4.9}$$

where X_L is the amplitude of the fundamental component of the error signal when limit cycling. Thus at limit cycle frequency, the imaginary part of the loop frequency response function is zero and by rationalizing the left-hand side of equation 4.9 and equating the imaginary part to zero we obtain

$$\frac{4}{\pi X_L \omega_L} \left\{ \frac{\omega_L \cos \omega_L t_d + \sin \omega_L t_d}{1 + \omega_L^2} \right\} = 1$$

or

$$x_L = X_L = \frac{2}{\pi^2} P \left\{ \frac{\omega_L \cos \omega_L t_d + \sin \omega_L t_d}{1 + \omega_L^2} \right\} \tag{4.10}$$

where x_L is the dimensionless measure of misalignment limit cycle amplitude and P is the time limit cycle period.

The values of P and x_L from equations 4.8 and 4.10 are plotted as shown

in Fig. 4.4. From Fig. 4.4 we can predict limit cycle amplitude and period in an on-off system.

Equation 4.9 can be rewritten as

$$G(j\omega) = -\frac{1}{G_D} \tag{4.11}$$

Equation 4.11 may be analyzed for stability by Nyquies method or by root locus method if the describing function G_D is amplitude-dependent only.

When G_D is amplitude- but not frequency-dependent, it is possible to reduce a very simple criterion for predicting stability or instability. The basis for the criterion is that the leftside of equation 4.11 is only frequency-dependent and the right side is only amplitude-dependent. Marginal stability as predicted from the Nyquist criterion and the describing function concept requires a solution of equation 4.11, or in other words, an intersection of these two curves. At the point of intersection, the frequency parameter on the one curve and the amplitude parameter on the other give the frequency and amplitude of the corresponding oscillation. If there is no intersection, then the system is presumably stable or unstable for all conditions of operation. Since it is difficult to find a rigorous definition of stability for nonlinear systems, the above statement must be regarded as provisional, pending more detailed investigation or more careful definition of stability for unusual systems.

From equation 4.11, we have $G(j\omega) = \dfrac{e^{-j\omega t_d}}{j\omega(j\omega+1)}$,

$$-\frac{1}{G_D} = -\frac{\pi X}{4Y} = -1.5708X$$

if we put $t_d = .2$ seconds and by assigning some values of ω, we can have the following data; and from these data the Nyquist plot can be plotted as shown in Fig. 4.5.

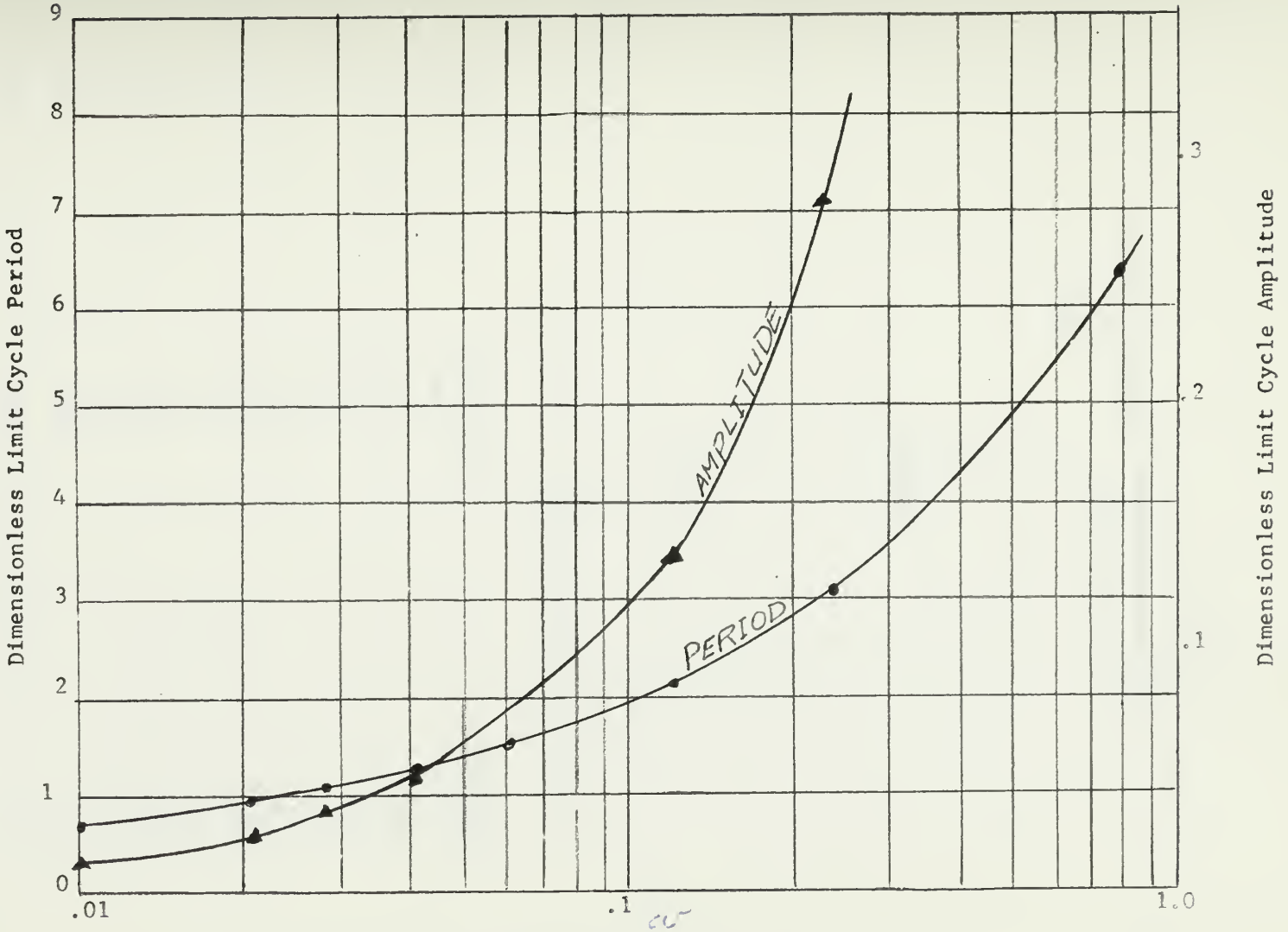


Fig. 4-4 Curves for predicting limit cycle amplitude and period in an on-off system

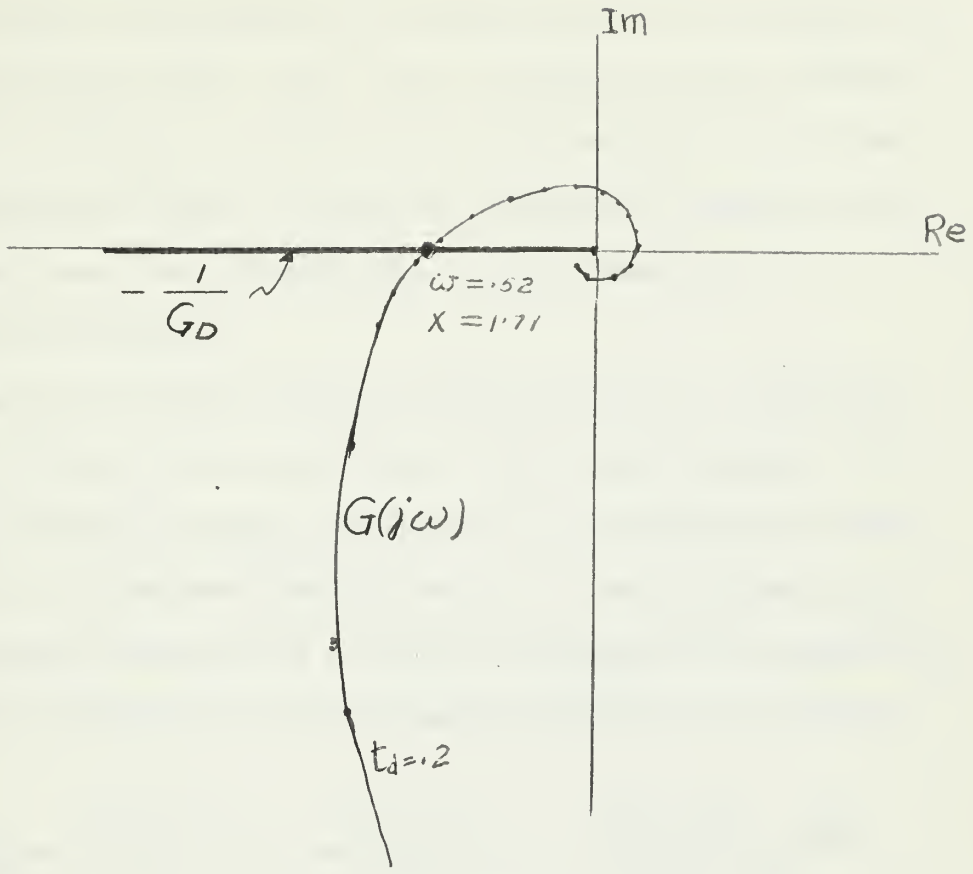


Fig. 4-5 Nyquist Plot of Eq. 4-11

ω	$\lvert G(j\omega)\rvert$	$\angle G(j\omega)$
0	∞	-90
.2	4.9	-124.2
.3	3.19	-147
.4	2.32	-167.6
.5	1.79	-173.9
.6	1.43	-189.6
.8	.973	-220.12
1.0	.707	-249.5
1.2	.641	-277.7
1.6	.529	-331.4
2.0	.447	-382.4

From the Nyquist plot, a limit cycle is observed; it is at a point where $\omega = .52$ and $X = 1.71$. (The intersection of these two curves). The intersection may indicate a stability boundary in the sense that excitation

at level below that corresponding to the intersection will be unstable for the system, and should reach a limit cycle in which the input to the nonlinear element is roughly sinusoidal with angular frequency $\omega = .52$ and X = 1.71. The system should be unstable for low level of excitation and stable for level above that corresponding to the intersection.

4.4.2 The root locus method

Because the describing function of an ideal relay is independent of frequency, so that the root locus method can be readily applied. The nonlinearity introduces no phase shift but only a gain dependent on amplitude. It will be assumed here that the behavior of the nonlinear system can be approximated through use of a root-locus diagram for the system with the nonlinearity eliminated. By block diagram manupulation as follows;

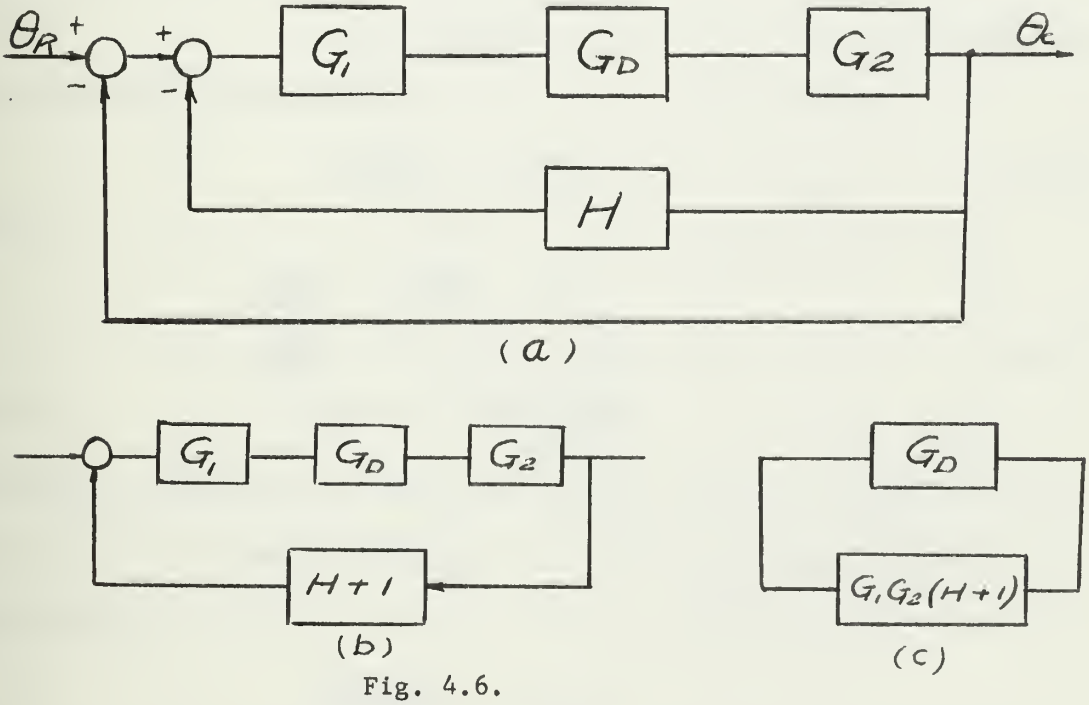


Fig. 4.6.

and algebraic manupulation provides the stability relationship, the characteristic equation

$$G_1 G_2 G_D(H+1) = -1 \qquad (4\text{-}12)$$

$$G(S) = -1 \tag{4.13}$$

$$G(s) = \frac{K'(K_t S + 1)\, e^{-st}}{S(S + 1)}$$

$$K' = \frac{4Y}{\pi X}$$

$$K_t S = H$$

For $t_d = .5$ second,

put $K_t = 1.5$ then we have

$$G(S) = \frac{K\,(S + 0.67)e^{-.5S}}{S\,(S + 1)}$$

Equation 4.13 is complex and may be split into two equations by equating the magnitude and phase angle on both sides respectively. Thus we have the magnitude equation

$$|G(S)| = 1 \tag{4.14}$$

and the phase angle equation

$$\angle G(s) = \pm 180° \tag{4.15}$$

where

$$\angle G(s) = \angle G'(s) - \omega T$$

$$\angle G'(s) - \omega T = \pm 180° \tag{4.16}$$

where $$G'(s) = \frac{K\,(s + 0.67)}{s\,(s + 1)}$$

a method is introduced to obtain the required locus of equation (4.16).[4] First it is to construct a family of root loci (or phase-angle loci) of the equation

$$\angle G'(s) = \phi_1$$

where ϕ_1 is a constant phase angle. Thus a family of loci can be plotted by various values of ϕ_1, Fig. 4.7 illustrates a family of root loci for the transfer function $G'(s)$. Since the locus with phase angle ϕ_1 is symmetrical

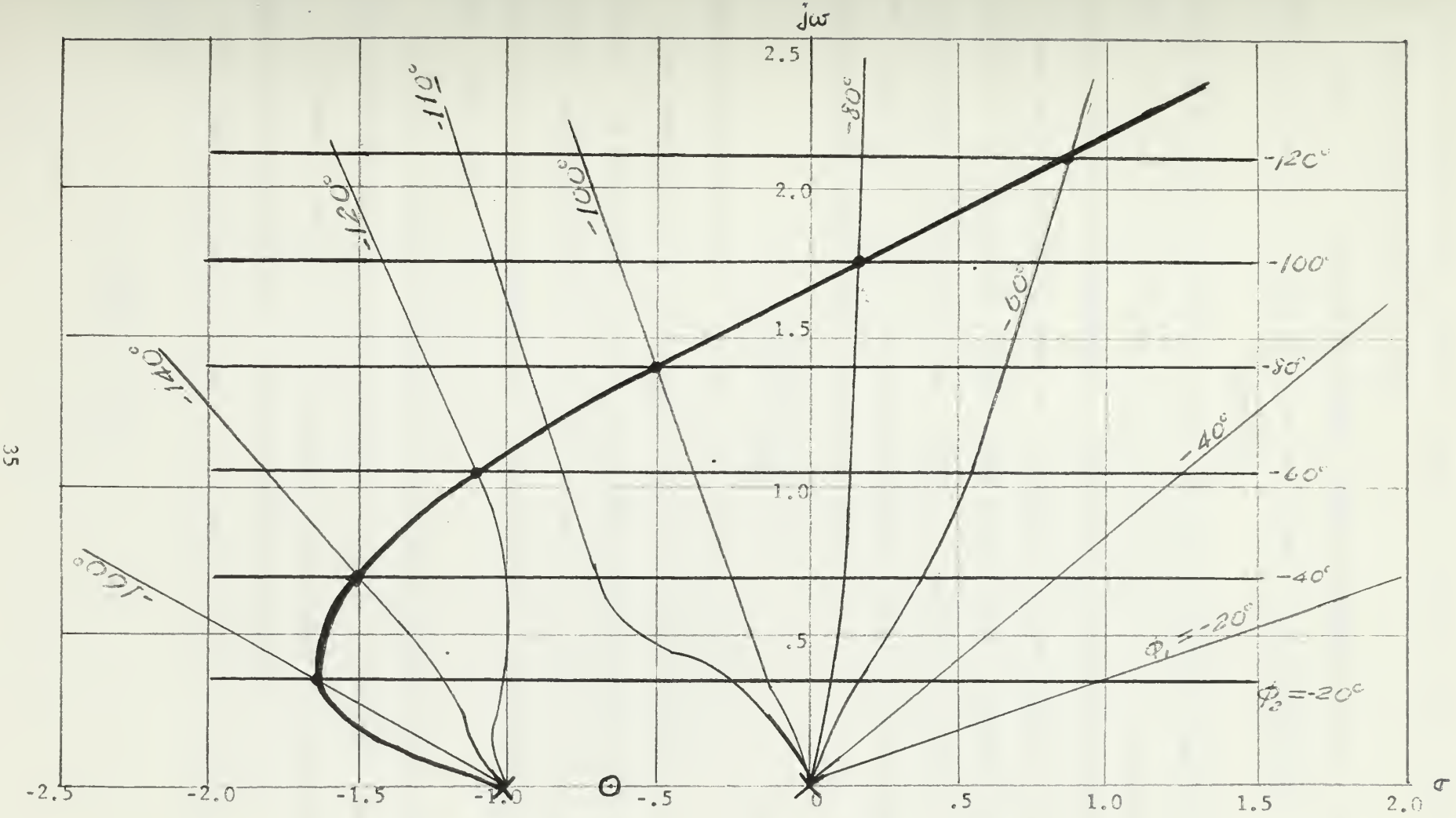


Fig. 4-7 Construction of the Required locus with Transportation lag

to that with phase angle (-ϕ) with respect to the real axis, only the loci on the upper half plane are shown.

Next, it is to construct another family of phase-angle loci (the name dead time lag loci is suggested) on the s-plane from the relation

$$-\omega T = \phi_2$$

the dead time lag loci family is also plotted in the same sheet as shown in Fig. 4.7. Then by superimposing these two families of loci on the same s-plane, the point of intersection of the corresponding curves with the sum of phase angles ϕ_1 and ϕ_2 equal to 180 degrees are the points of the locus required. The curve drawn through these points is the required locus. It is the required locus because each point on it satisfies equation 4.16. The determination of the gain can be carried out then by the criterion in the root-locus method. The magnitude equation is

$$K = \frac{1}{|G(\sigma + j\omega)|} e^{-Ts} \qquad (4.17)$$

which will give the gain K of any point on the required locus.

The required roots are the points where the value of gain K is equal to one. The gain limit, which is the maximum gain below which the system can have damped oscillation, can be found readily. This is the value of gain at the point of intersection of the required locus and the imaginary axis. This value is found to be 1.21.

This given function transfer has an infinite number of roots, which can be obtained by constructing additional branches of required locus.

In general, the number of influential roots depends on the relative magnitudes of the dead time T and the other time constant in G'(s).

The limit cycle occurs at ω = 1.67 rad/sec. and k = 1.21 where $k = k'k_t$

$$K' = \frac{4Y}{\pi X} = \frac{2}{\pi X}$$

$$k_t = 1.5$$

$$K = 1.21 = \frac{2}{\pi X} \quad K_t = \frac{2}{\pi X} \times 1.5$$

$$X = \frac{2 \times 1.5}{\pi \times 1.21} = .788$$

5. System with Dead Time.

5.1. Transfer function associated with pure dead time

If the input to component with a pure dead time characteristic is f(t) u(t), then the output of the component is f(t - T), where $T > 0$ is the dead time.[7] The transfer function associated with pure dead time is, therefore,

$$G(s) = \frac{e^{-TS}F(s)}{F(s)} \tag{5.1}$$

$$= e^{-TS} \tag{5.2}$$

For sinusoidal input at frequency ω, the transfer function given in equation 5.2 becomes

$$G(j\omega) = e^{-j\omega T} \tag{5.3}$$

The gain is thus seen to be independent of the frequency and equal to unity, and the phase shift is seen to be directly proportional to frequency.

5.2 Frequency response of a system with dead time.

It is of interest to compare the frequency characteristics of a unit with pure dead time T and a unit with a first-order lag transfer function with time constant equal to T. That is,

$$G_1(s) = e^{-TS} \tag{5.4}$$

and

$$G_2(S) = \frac{1}{TS + 1} \tag{5.5}$$

The gain associated with $G_1(j\omega)$ and $G_2(j\omega)$ are

$$|G_1(j\omega)| = 1 \tag{5.6}$$

$$|G_2(j\omega)| = \frac{1}{\sqrt{1 + \omega^2 T^2}} \tag{5.7}$$

and their respective phase angles are

$$\phi_1 = -\omega T \tag{5.8}$$

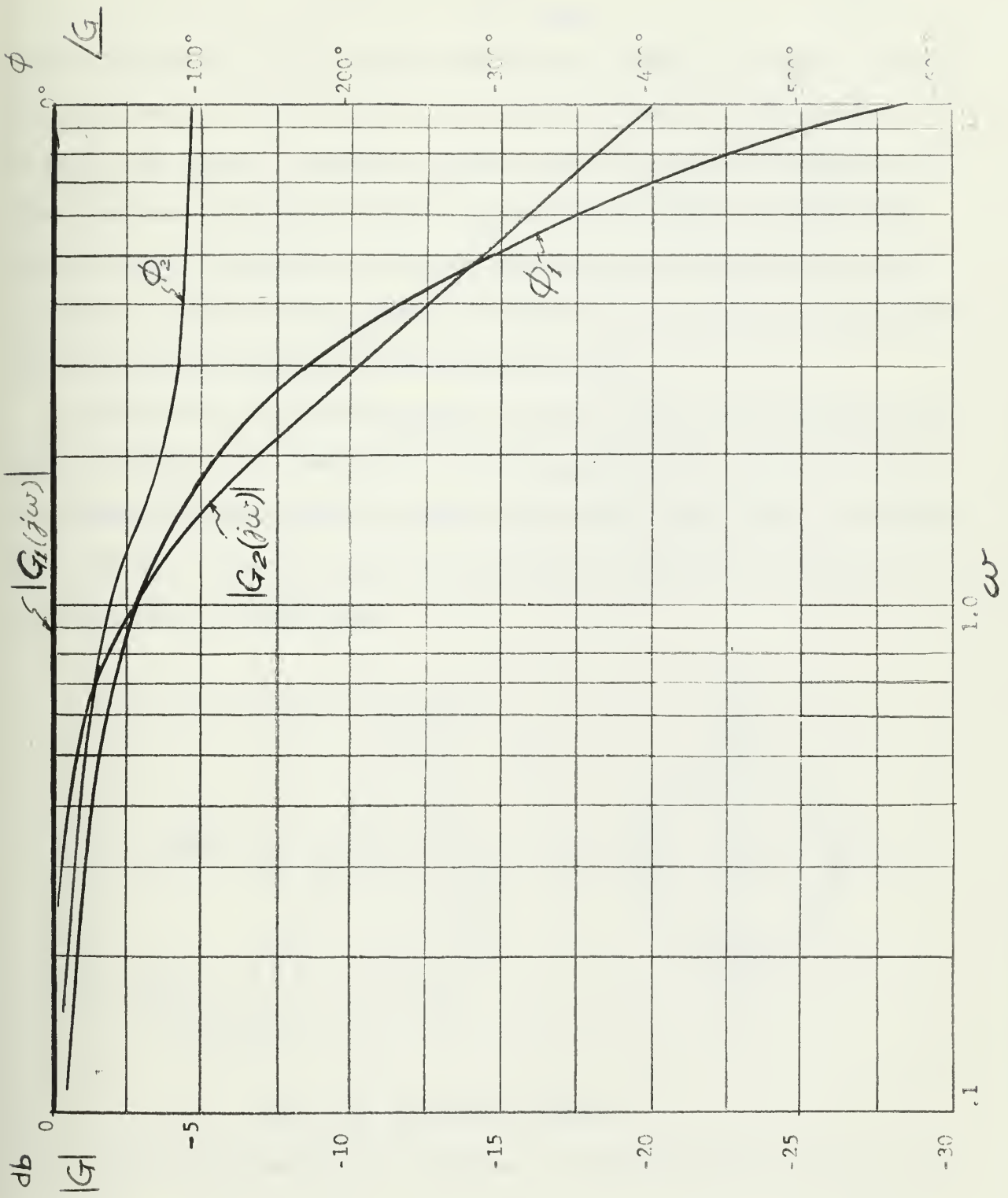


Fig. 5-1 Bode diagram for a pure dead-time unit and a first-order lag unit with time constant of the latter equal to the dead time of the f

and

$$\phi_2 = - \arctan \omega T \qquad (5.9)$$

The Bode diagram for the two units are plotted as shown in Fig. 5.1, where it can be seen that for $\omega T \ll 1$, the units have nearly the same effect on a given input signal. Generally, however, since for higher frequencies the dead-time unit exhibits more phase lag and more gain than the first-order lag unit, the introduction of the dead-time unit in cascade within a given closed-loop system has more adverse an effect on the stability of the system than does the introduction of the first-order unit.

Dead time has the characteristic of a pure phase shift since the magnitude is always one. Seen easily from the graphs drawn in Fig. 5.2 that no other dynamic element builds up phase lag so much as dead time. In combination with other dynamic elements the phase lag of dead time is simply added to obtain the total phase lag.

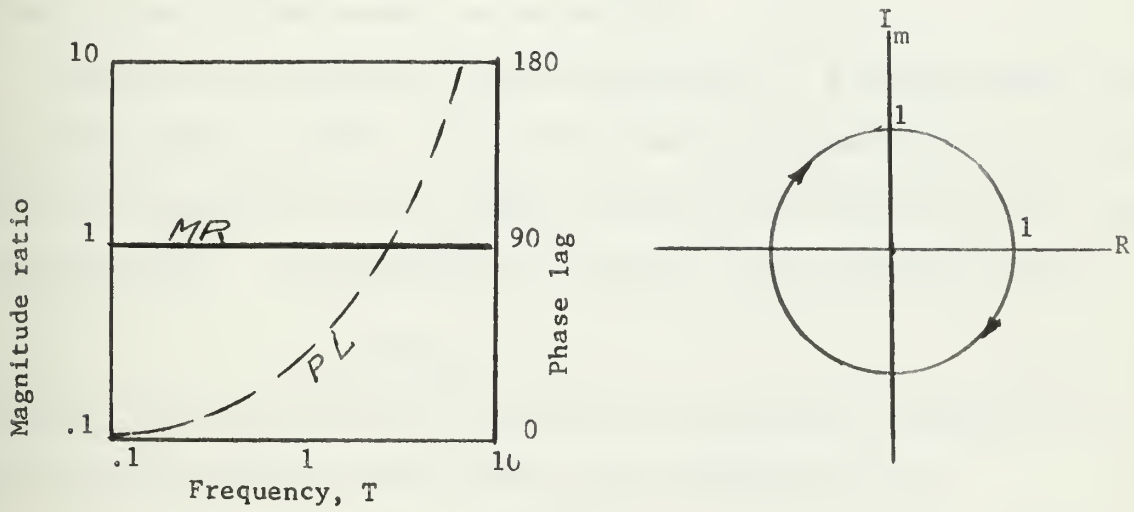


Fig. 5-2. Dead-time element.

A useful approximation for dead time T is the following function from the Pade list,

$$G(S) = \frac{2 - TS}{2 + TS} \qquad (5.10)$$

The magnitude of this function is everywhere 1.0 and the phase lag corresponds very closely to dead time up to frequencies given by $\omega T = 0.5$. In addition, this function may be easier to incorporate into complex system analysis.

The destabilizing effect of the transportation lag, as dead time is also known, and clearly illustrated in a Nyquist diagram. For example, if a unit with a pure dead-time characteristic given by

$$G_1(S) = e^{-TS} \tag{5.11}$$

is connected in cascade with a component with transfer function

$$G_2(S) = \frac{2}{S(S + 1)} \tag{5.12}$$

as in Fig. 5.3, the Nyquist diagram for the system for various values of the dead-time appears as in Fig. 5.4. As can be seen in this graph, for small values of transportation lag the introduction of the dead-time unit merely results in a decrease in relative stability, but for large values of transportation lag the system can be unstable.

The effect upon stability of the introduction of a transportation lag in cascade within a closed-loop system of the type illustrated in Fig. 5-5 (a) can be determined from the Nyquist curve for the original system in the following manner. The characteristic equation of the original system is

$$1 + G_a(s)G_b(s) = 0 \tag{5.13}$$

and that of the system which is obtained by introducing a transportation lag T within either the forward path or the feedback path is

$$1 + G_a(s)G_b(s)e^{-Ts} = 0 \tag{5.14}$$

Equation (5.14) can be rewritten as

$$G_a(s)G_b(s) = -e^{+Ts} \tag{5.15}$$

or, in terms of frequency response, as

$$G_a(j\omega)G_b(j\omega) = -e^{j\omega T} \tag{5.16}$$

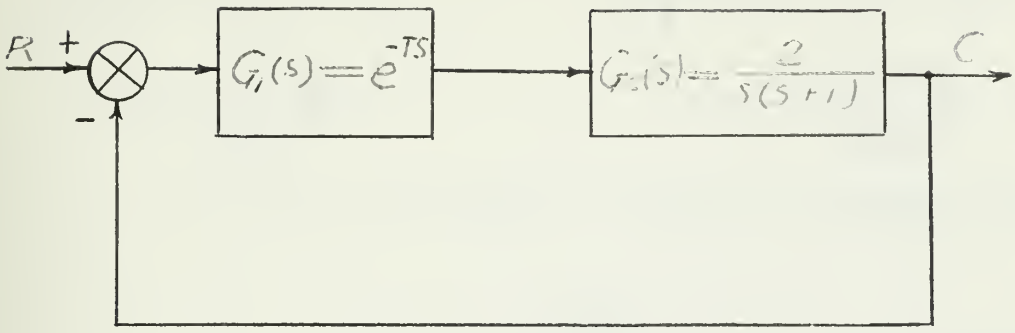


Fig. 5-3. Block diagram of a simple feedback system with transportation lag.

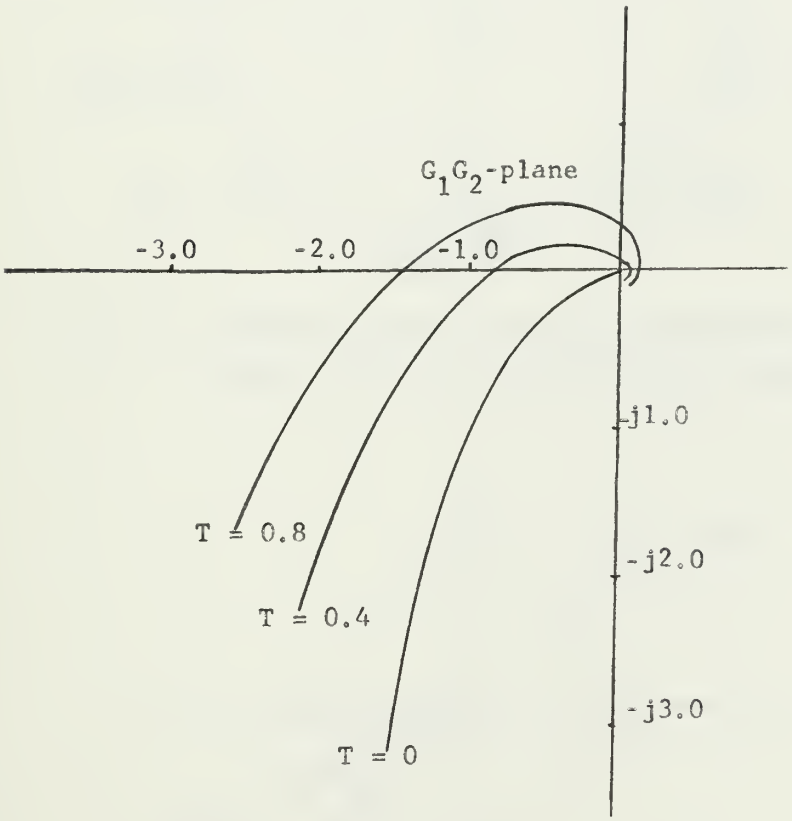


Fig. 5-4. Nyquist diagram for the system illustrated in Fig. 5-3.

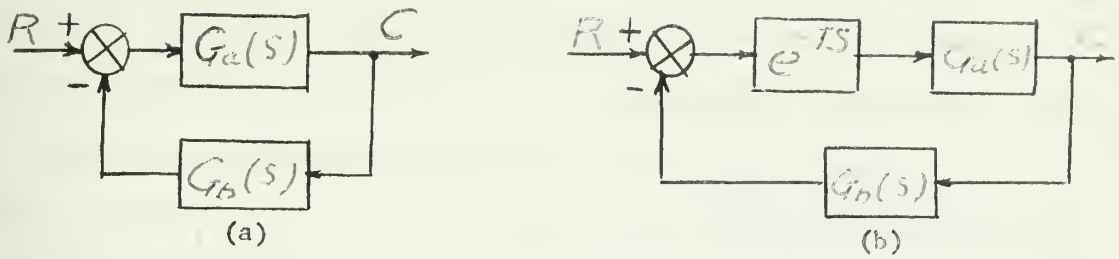


Fig. 5-5 Block diagram of a single loop feedback system

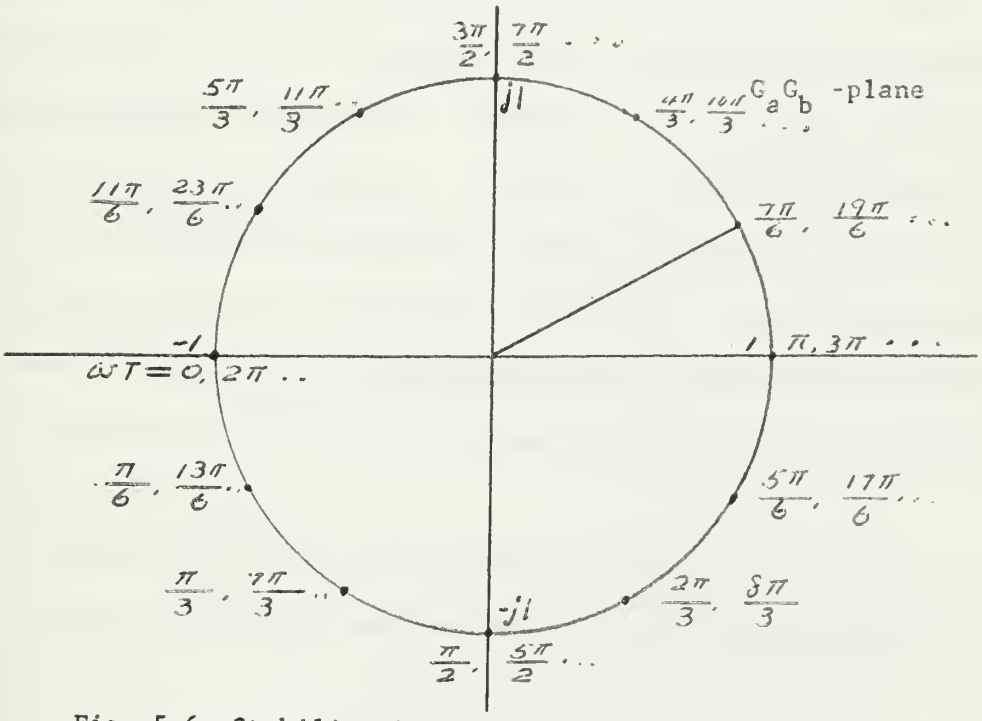


Fig. 5-6 Stability locus in the G_aG_b -plane for a system of the type shown in Fig. 5-5(b).

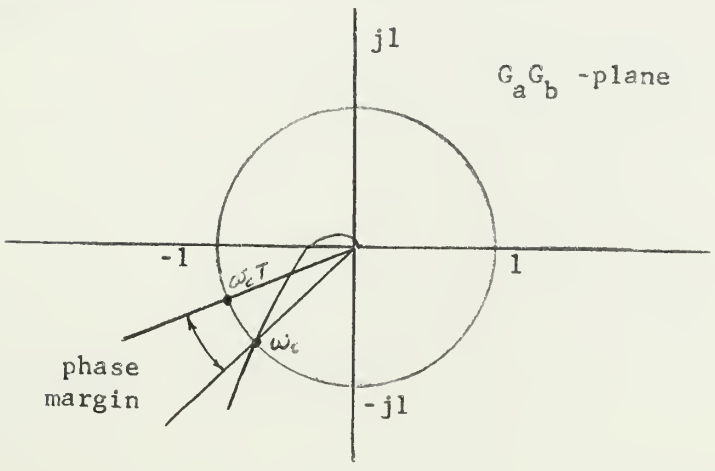


Fig. 5-7 Determination of phase margin in a system with transportation lag.

The locus of $G_a(j\omega)G_b(j\omega)$ is the Nyquist curve for the original system; equation (5-16) states that the effect of introducing transportation lag T is to shift the critical stability point in the G_aG_b -plane from -1 + j0 to $-\cos\omega T - j\sin\omega T$. That is, the critical point becomes a critical locus, in the form of the unit cycle in the G_aG_b -plane, as shown in Fig. 5-6.

The characteristic equation (5-16) is satisfied only if the locus of $G_a(j\omega)G_b(j\omega)$ intersects the stability locus in such a manner that the frequency on the Nyquist curve and the frequency on the stability locus are the same at the point of intersection. If the two frequencies are not equal at the point of intersection, a measure of relative stability is afforded by the stability locus. Thus, if the Nyquist diagram for a given system has the appearance shown in Fig. 5-7, the phase margin can be determined as follows:

1. Determine the cross-over frequency ω_c , at whichthe Nyquist curve crosses the unit circle.

2. Locate the point on the stability locus for $\omega = \omega_c$.

3. Draw a radial line segment from the origin through each of these points.

4. Measure the angle between these line segments. This angle is the phase margin, with the same positive sense as in the case of a system without transportation lag.

6. Transportation lag.

6.1. General

The phenomenon of dead time, sometimes called transportation lag or time delay or ideal time lag, can best be categorized as a lack of change in output for a specified interval of time following a change in input, this is illustrated in Fig. 6.1.

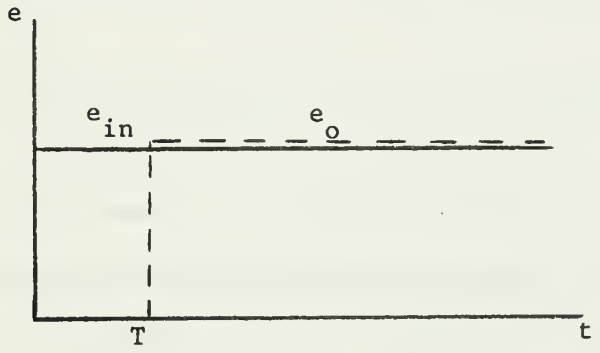


Fig. 6.1

Transportation lag may be encountered in process involving pressure, temperature, and also may be found in process involving chemical reactions when a finite time must elapse before a reaction begins to occur.

Occasionally in analyzing engineering systems, it becomes necessary to simulate time delay (or transportation lag). For example in studying transmission processes through hydraulic or pneumatic lines or in studying the motion of radio active particles in neuclear reactors it is necessary to produce an output voltage which is identical with an input voltage but displaced in time by a specified amount.

Transportation lag may be expressed by the following mathematical relationship: Given $f(t)$ for $t \geqq 0$, the function $f(t - T)$ differs from the function $f(t)$ by the dead time T, where $f(t - T)$ is defined to be zero for $t < T$.

A direct and accurate method for generating such time delay is to employ a magnetic tape recorder in which the space between the recording and reading heads is adjustable, by suitably controlling the relative location of the two heads and the speed of the tape a wide range of time delay can be simulated.

Approximate methods which do not require special auxiliary equipment, once an analog computer is available, are based upon the Laplace-transformation relationship

$$\mathcal{L}[f(t-T)] = F(s)e^{-sT} = \mathcal{L}[f(t)]e^{-sT} \qquad (6.1)$$

since e^{-sT} cannot be expressed as a rational function, the transfer function for dead time can be simulated only approximately, thus the remaining methods depend on an approximate series or technique of network synthesis.[9]

The Taylor series expansion of this function is

$$e^{-sT} = 1 - TS + \frac{(TS)^2}{2!} - \frac{(TS)^3}{3!} + \frac{(TS)^4}{4!} - \cdots + \frac{(TS)^n}{n!} \qquad (6.2)$$

The generation of the terms of this series would require successive differentiation with attendent noise problems. The rate of convergence of equation 6.2 is further relatively slow, particularly for large values of T, or high frequencies. For this reason the Taylor series is not well suited to the generation of transportation lag for problems involving high frequencies or long time constants.

6.2 Pade approximation

A more practical approach to the generation of the transportation lag function is to employ the so called Pade' approximation.[9] From equation 6.1 the transfer function of transportation lag is

$$e^{-sT} = \frac{\mathcal{L}[f(t-T)]}{\mathcal{L}[f(t)]} = \frac{E_o}{E_i} \qquad (6.3)$$

substituting $s = j\omega$ to find the frequency response gives

$$\frac{E_o}{E_i} = e^{-j\omega T} = \cos\omega T - j\sin\omega T$$

thus

$$\left|\frac{E_o}{E_i}\right| = \sqrt{\cos^2\omega T + \sin^2\omega T} = 1 \tag{6.4}$$

$$\phi = \tan^{-1}\left(-\frac{\sin\omega T}{\cos\omega T}\right) = -\omega T \tag{6.5}$$

from the above equations the frequency transfer function of the transportation lag is characterized by an amplitude ratio which is unity for all frequencies and a phase angle which varies linearly with frequency over the desired range of interest. From equations 6.4 and 6.5 it is seen that the quotient of any pair of complex-conjugate polynominals has an amplitude of unity and also possess considerable phase shift.

In the Pade' approximation the exponential is approximated as the ratio of two polynominals in s, according to

$$e^{-sT} = \frac{P_n(s)}{Q_m(s)} \tag{6.6}$$

where subscripts m and n refer to highest power of s in the numerator and denominator, respectively. Expressed in its basic form, the Pade' approximation is

$$e^x = \lim_{(u+v)\to\infty} \frac{F_{u,v}(x)}{G_{u,v}(x)} \tag{6.7}$$

where

$$F_{u,v}(x) = 1 + \frac{vx}{(u+v)!} + \frac{v(v-1)x^2}{(u+v)(u+v-1)2!} + \cdots + \frac{v(v-1)(v-2)\cdots 2\cdot 1\, x^v}{(u+v)(u+v-1)\cdots(u+1)v!} \tag{6.8}$$

$$G_{u,v}(x) = 1 - \frac{ux}{(v+u)!} + \frac{u(u-1)x^2}{(v+u)(v+u-1)2!} + \cdots + (-1)^u \frac{u(u-1)(u-2)\cdots 2\cdot 1\, x^2}{(v+u)(v+u-1)\cdots(v+1)u!} \tag{6.9}$$

the convergence for this series expansion is quite rapid and often values of u and v of 2 will give good accuracy for short time lag and low frequencies.

The Pade' approximation of e^{-sT} for u = v = 2 and if x = -sT is

$$e^{-ST} \approx \frac{1 - \frac{2}{4}(ST) + \frac{2}{4!}(ST)^2}{1 + \frac{2}{4}(ST) + \frac{2}{4!}(ST)^2} \tag{6.10}$$

$$e^{-ST} \approx \frac{S^2T^2 - 6ST + 12}{S^2T^2 + 6ST + 12} \tag{6.11}$$

the higher the powers of the numerator and denominator, the more terms of the series can be matched and the smaller the error in the time delay function. Equation 6.11 is the second order Pade' approximation. The fourth order Pade' approximation (u = v = 4) for e^{-sT} is

$$e^{-ST} = \frac{(ST)^4 - 20(ST)^3 + 180(ST)^2 - 840(ST) + 1680}{(ST)^4 + 20(ST)^3 + 180(ST)^2 + 840(ST) + 1680} \tag{6.12}$$

Higher order pade' approximations provide improved accuracy and frequency response. It is noted that the order of the characteristic equation is raised by the order of the approximant used to simulate e^{-sT}. If the second order approximant is used, two additional roots are introduced into the characteristic equation.

6.3 Simulation of transportation lag

Equation 5.12 can be rewritten as

$$\frac{\mathcal{L}[f(t-T)]}{\mathcal{L}[f(t)]} = e^{-ST} = \frac{1 - \frac{20}{ST} + \frac{180}{(ST)^2} - \frac{840}{(ST)^3} + \frac{1680}{(ST)^4}}{1 + \frac{20}{ST} + \frac{180}{(ST)^2} + \frac{840}{(ST)^3} + \frac{1680}{(ST)^4}} \tag{6.13}$$

since the differential operator, D, and Laplace operator are interchangeable

when initial conditions are zero, equation 6.13 can be expressed as

$$\frac{f(t-T)}{f(t)} = \frac{1 - \frac{20}{DT} + \frac{180}{(DT)^2} - \frac{840}{(DT)^3} + \frac{1680}{(DT)^4}}{1 + \frac{20}{DT} + \frac{180}{(DT)^2} + \frac{840}{(DT)^3} + \frac{1680}{(DT)^4}} \qquad (6.14)$$

rearrange equation 6.14 and we get

$$\begin{aligned} f(t-T) = f(t) &- \left[\frac{20}{T} f(t) + \frac{20}{T} f(t-T)\right]\frac{1}{D} \\ &+ \left[\frac{180}{T^2} f(t) - \frac{180}{T^2} f(t-T)\right]\frac{1}{D^2} \\ &- \left[\frac{840}{T^3} f(t) + \frac{840}{T^3} f(t-T)\right]\frac{1}{D^3} \\ &+ \left[\frac{1680}{T^4} f(t) - \frac{1680}{T^4} f(t-T)\right]\frac{1}{D^4} \qquad (6.15) \end{aligned}$$

according to equation 6.15, the analog computer circuit can be simulated as shown in Fig. 6.2.

The time delay T in second is determined by the magnitude of the input resistors to the four integrators. Such a circuit works well only in a limited frequency range (direct current to approximately 50 cps). Instability results if the variation of the input signal f(t) is too rapid. It is therefore not feasible to apply a step function to this circuit. In case f(t) is not equal to zero at the commencement of the computer run, it is necessary to apply appropriate initial conditions to the integrators. The transportation lag devices described thus far are limited to applications in which the delay time T remains constant throughout the computer run.

The method of simulating transportation lag by the Pade' approximation yields what may be termed optimally flat phase characteristics at the origin but does not fit the phase characteristics as well over an extended frequency range. The major drawback to the use of the Pade' approximation to simulate transportation lag is their response to a step input or other

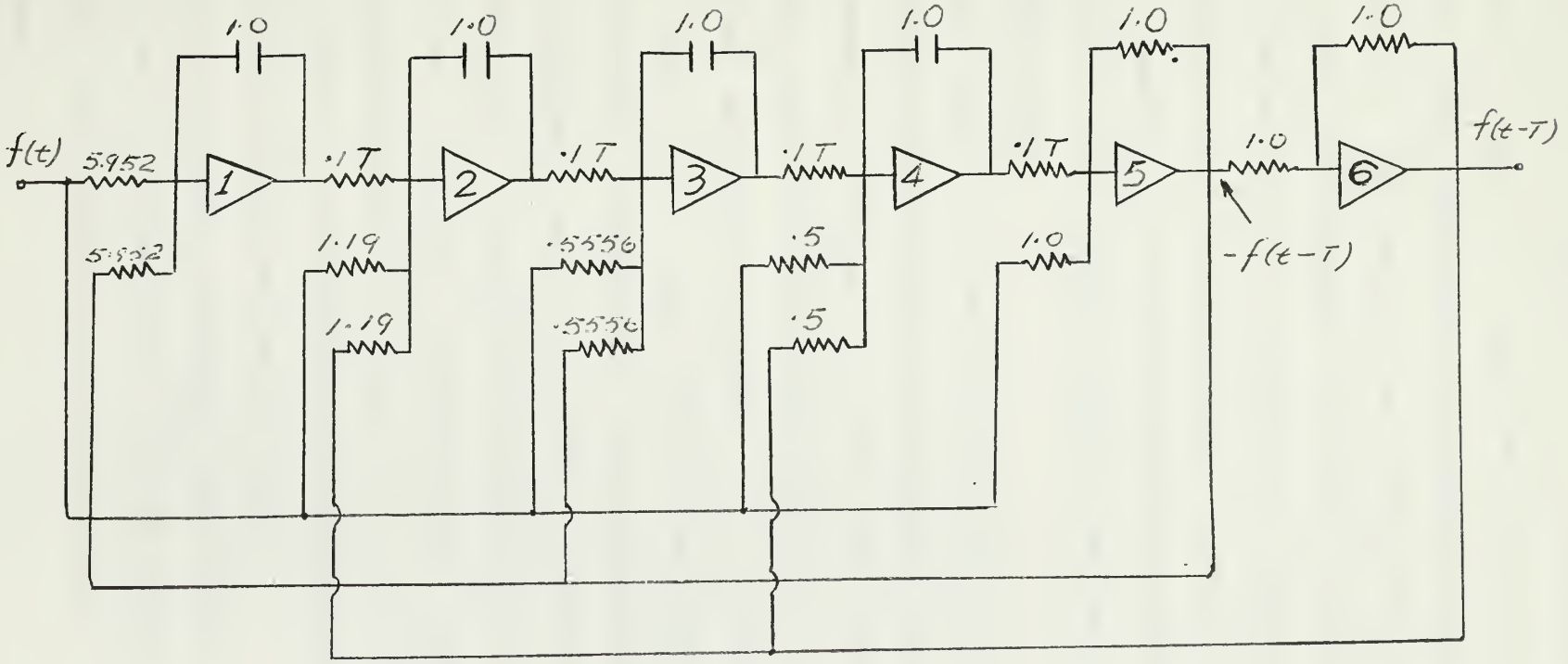


Fig. 6-2 Analog computer circuit for fourth order Pade approximation.

input having high frequency components, since the amplitude ratio is always unity but the phase shift is approximated only for low ωT products. The phase shift characteristics of certain of the Pade' approximation are shown in Fig. 6.3. As a result of the inability of this approximation to linearly phase shift high frequencies, the output of an approximant rings badly for high-frequency inputs. This behavior is illustrated in Fig. 6.4.

The actual location of the transportation lag transfer function in the forward path is immaterial and it may be placed at the end of the path in the computer where the high frequency components have been attenuated to the greatest possible extent.

In the great majority of cases the fourth order Pade' approximant is more than adequate for the simulation of transportation lag. In cases where linear phase shift of higher ωT products than those shown in Fig. 6.5 is desirable, higher order Pade' approximants or more complex curve fitting techniques may be used to simulate e^{-sT}. The phase shift characteristics of the sixth order and eighth order Pade' approximants are shown in Fig. 6.5. On the basis of this result it may be inferred that the order of the Pade' approximation may be increased to increase the magnitude of the ωT product that is linearly phase shifted without distorting the phase shift of low ωT products. Because the Pade' approximation may be expressed in a general form from which any order approximant may be developed, it is felt that the use of curve fitting techniques requires a needless expenditure of effort without any tangible benefit.

6.4 Generation of time delay from magnetic tape recorder.

The transportation lag generated by using the magnetic tape recorder is very accurate, and it saves a lot of time and labor for simulation by using amplifiers.

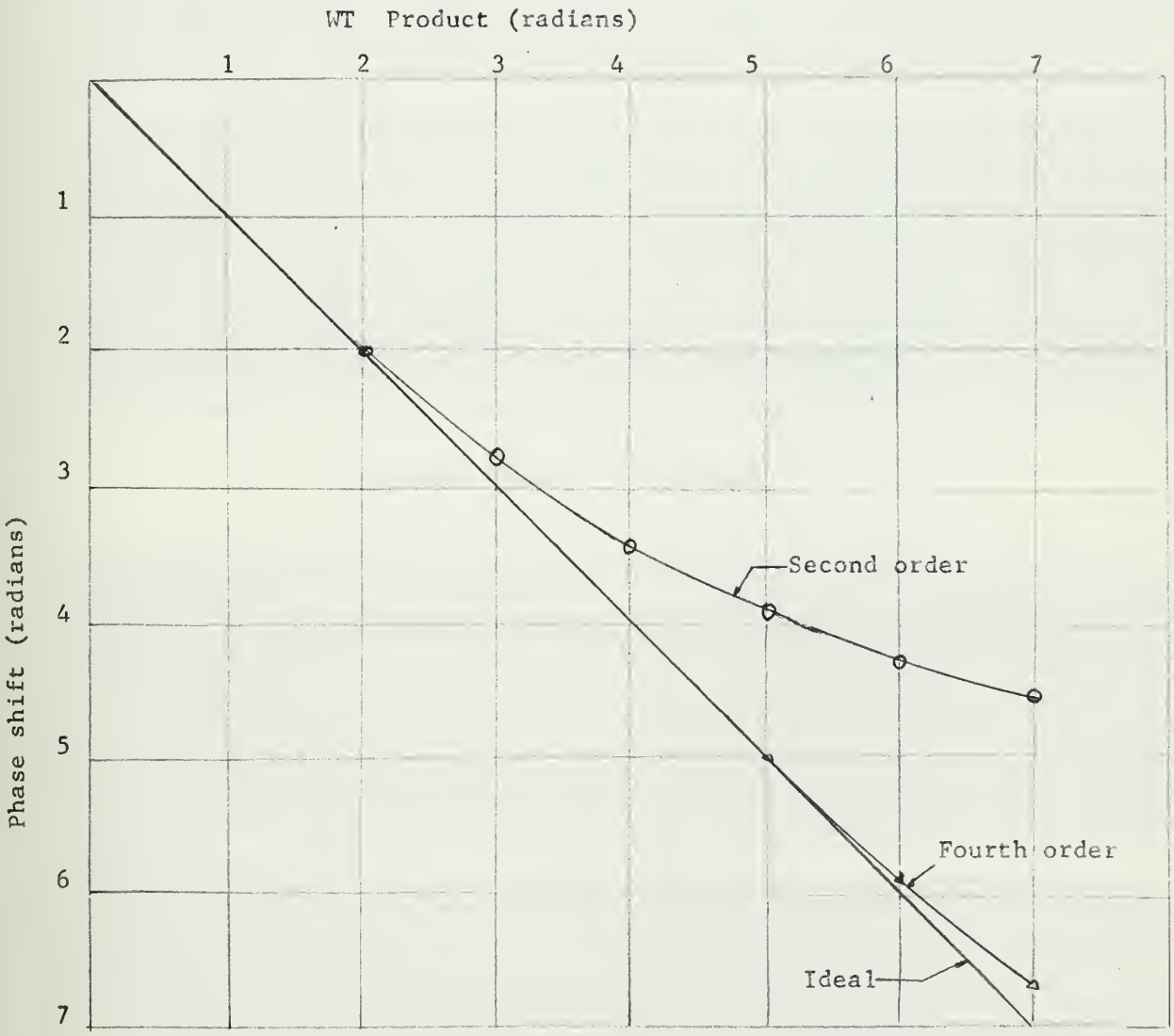
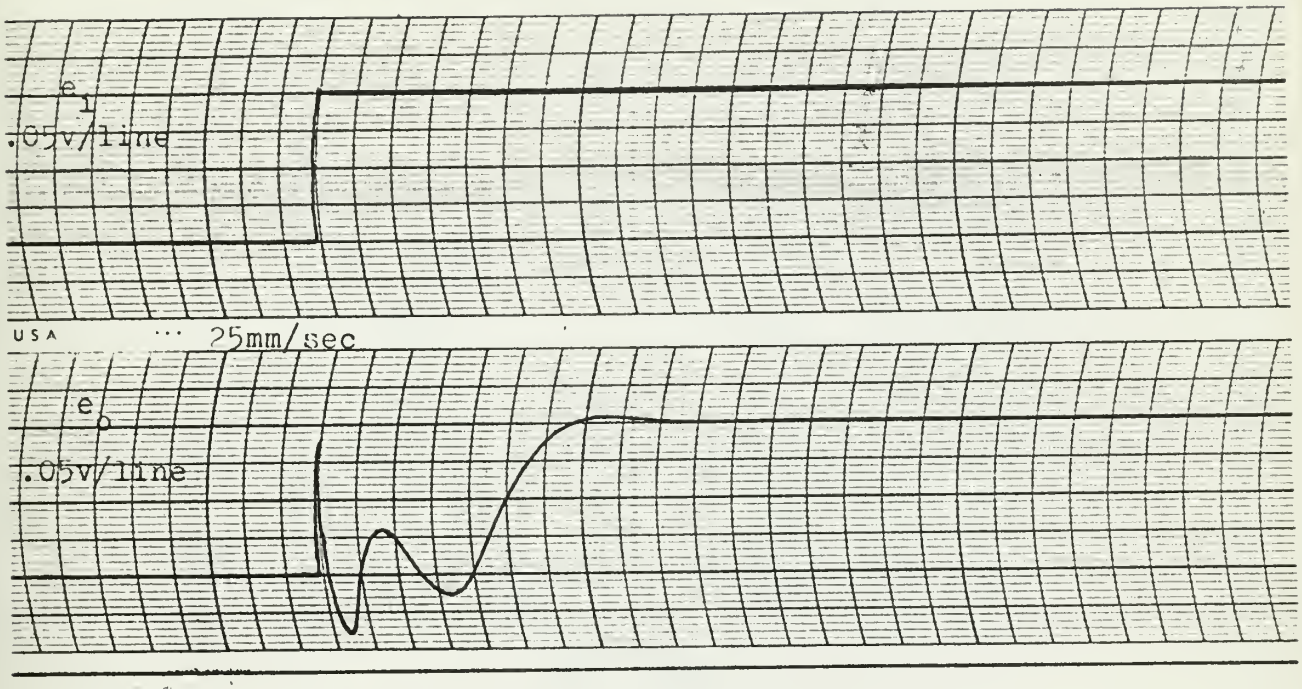
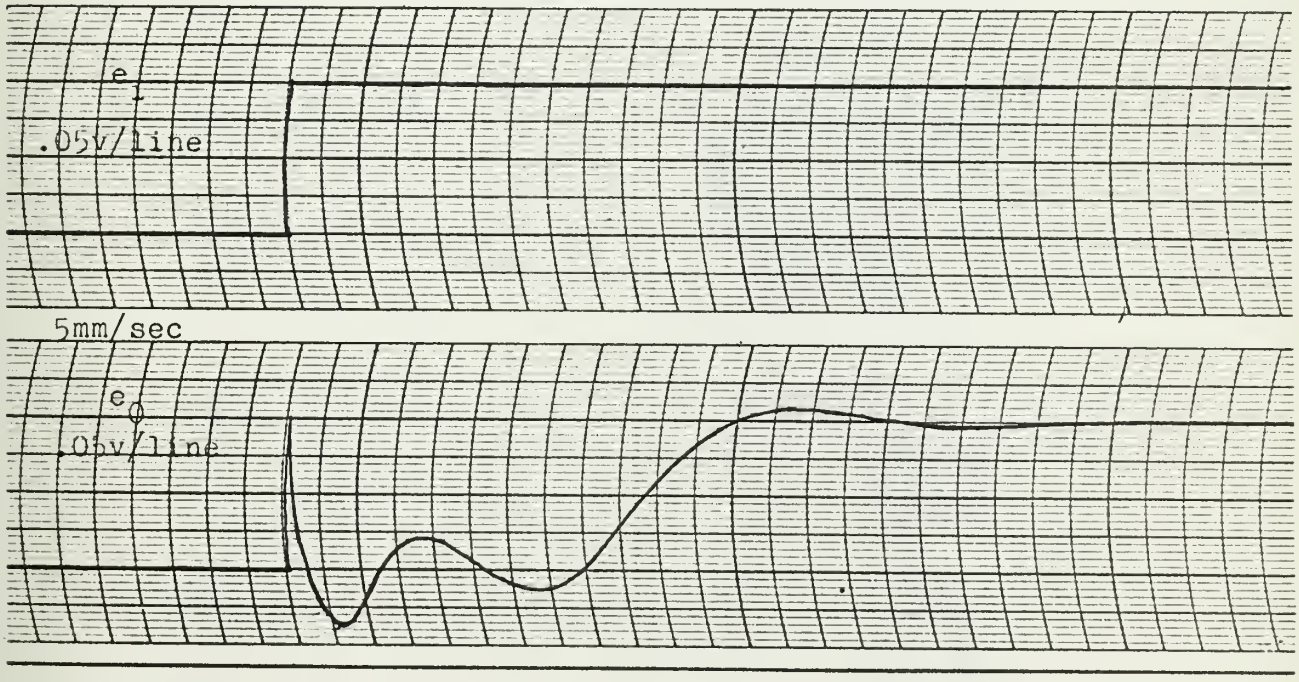


Fig. 6-3 Phase Shift Characteristics of the Pade Approximation



(a) delay time = 1.0 second



(b) delay time = 10 seconds

Fig. 6-4 Curves obtained from Pade approximation

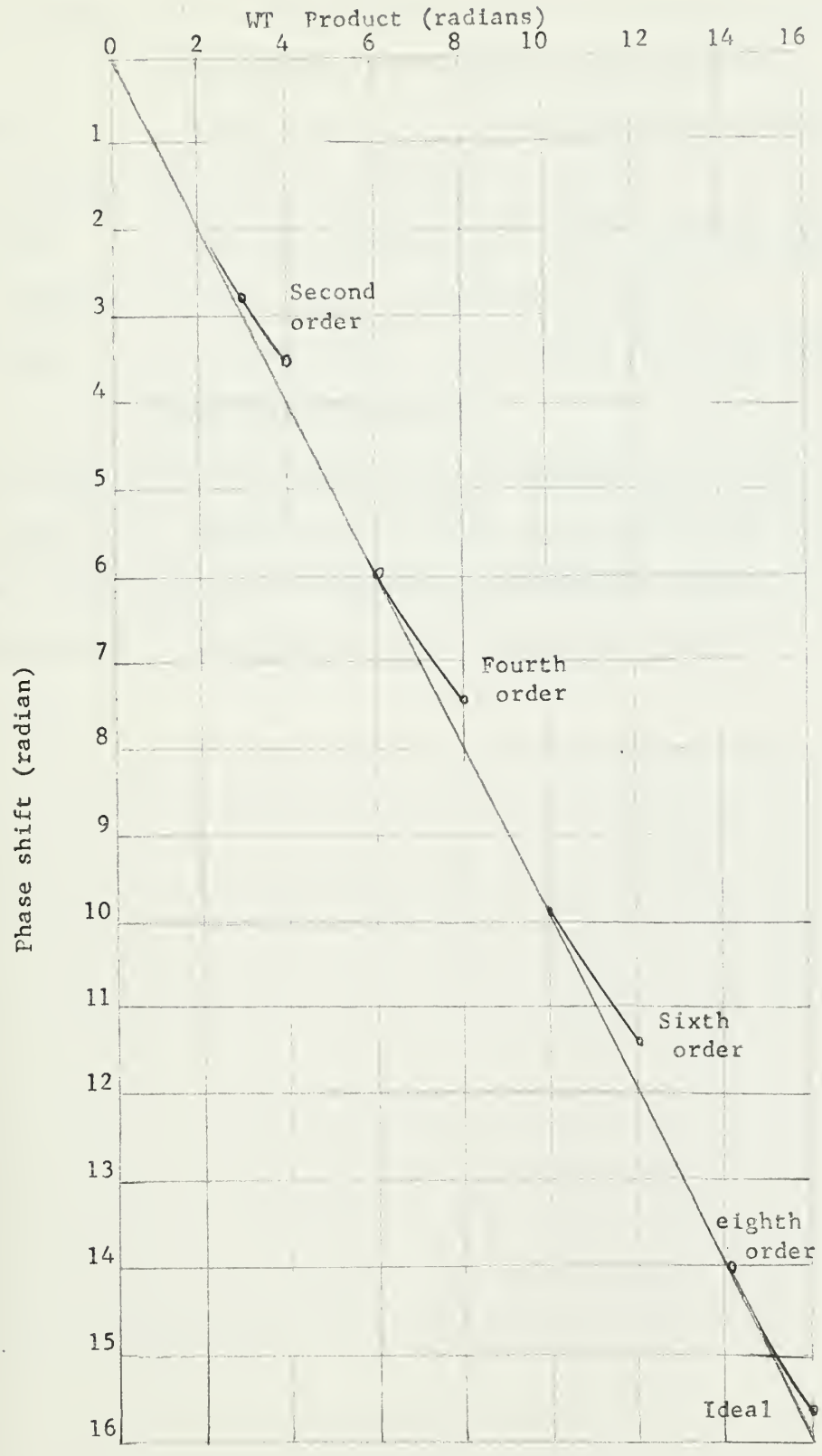


Fig. 6-5 Phase shif' Characteristics of the Pade Approximation

The magnetic tape recorder used has the following specifications:

Ampex Model FR1107 magnetic tape recorder

Tape speed: 3-3/4, 7-1/2, 15, 30, and 60 inches per second.

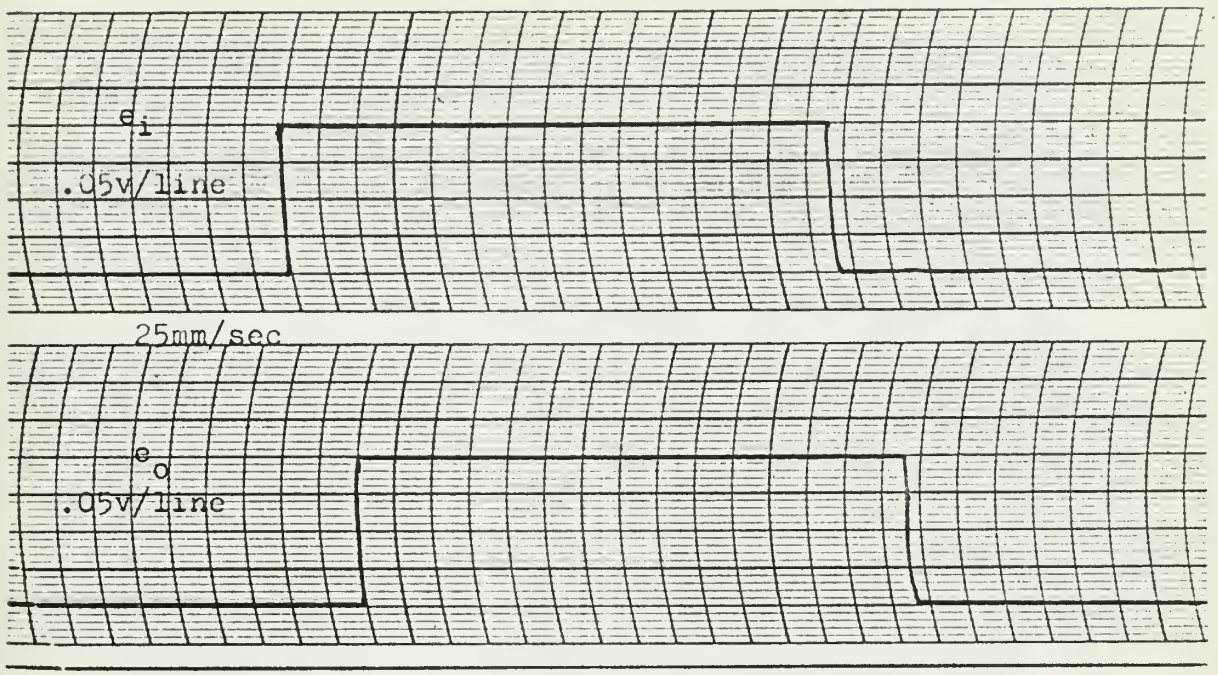
Input impedance: 100,000 ohms unbalanced to ground.

Input level: 1 volt rms recommended for normal recording level.

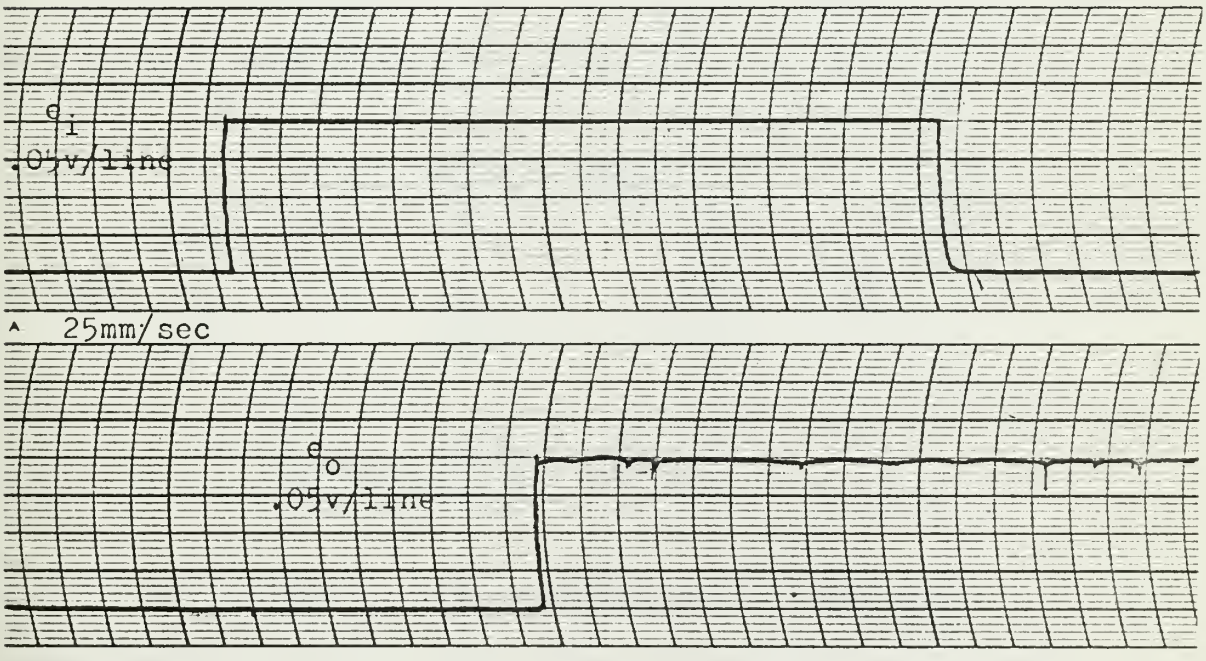
Output impedance: 125 ohms 75 μf, unbalanced to ground.

Output level: 1 volt rms normal, across 10,000 ohms load impedance at normal recording level.

The response curves obtained from the magnetic tape recorder are as shown in Fig. 6.6. Compare this with that obtained from analog computer as shown in Fig. 6.4, it is obvious that the magnetic tape recorder is most desirable for simulating the time delay, since it gives a more accurate result.



(a) delay time = .4 second



(b) delay time = 1.7 seconds

Fig.6-6 Curves obtained from magnetic tape recorder

7. System simulation.

7.1. Analog computer simulation

The block diagram of the system to be investigated is as shown in Fig. 7.1.

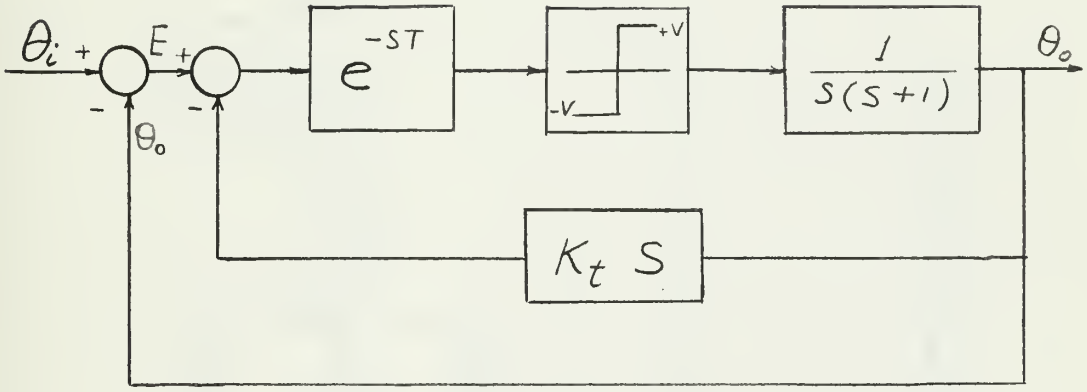


Fig. 7.1. The block diagram of the relay servo system.

The simulation of the transportation lag is shown in Fig. 6.2. The simulation of the ideal relay is as shown in Fig. 7.2.

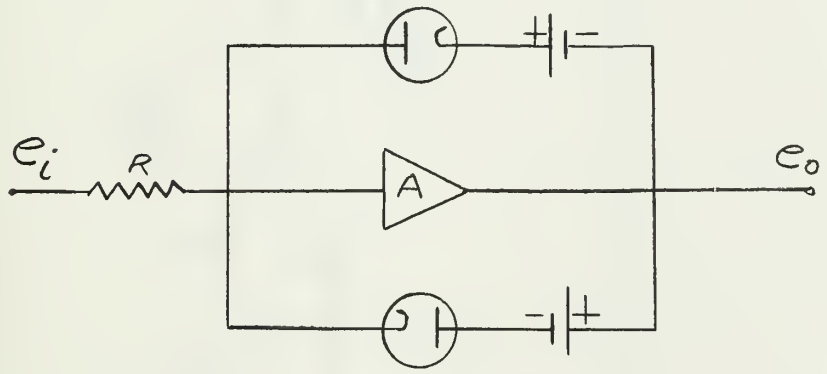


Fig. 7.2. The simulation of the ideal relay.

From the block diagram of the system, the analog computer circuit diagram can be simulated as shown in Fig. 7.3.

Because of its convenience and accuracy, the magnetic tape recorder is selected for the generation of transportation lag (e^{-sT}) through the experiments.

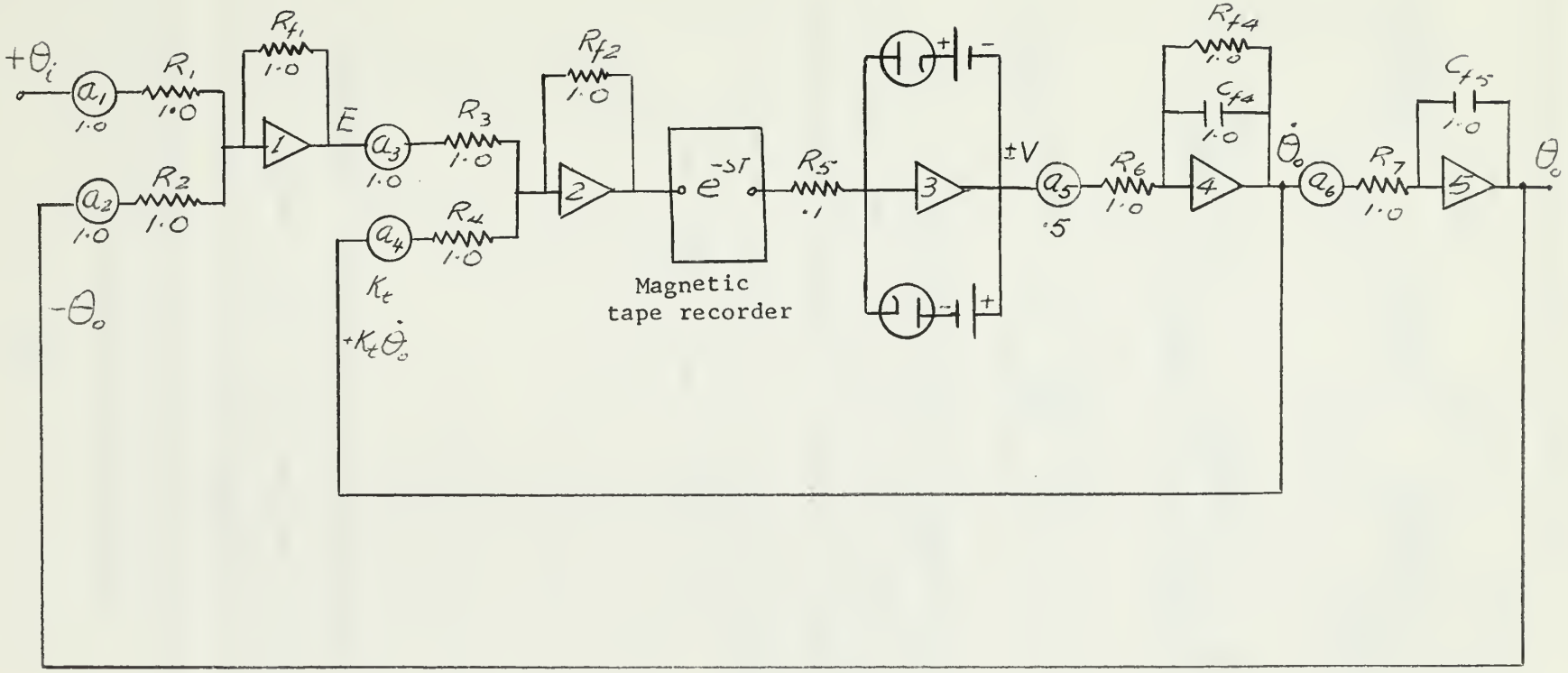


Fig. 7-3 Operational Diagram of Analog Computer setup for the block diagram of the relay servo system of Fig. 7-1.

7.2. Digital computer program

The block diagram of the system is shown below.

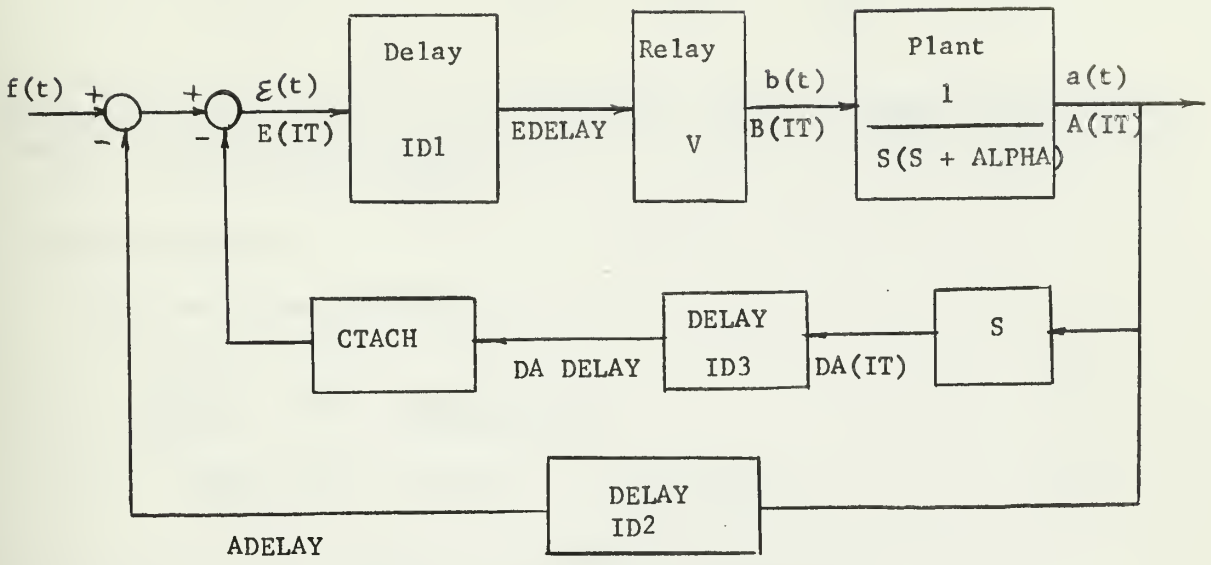
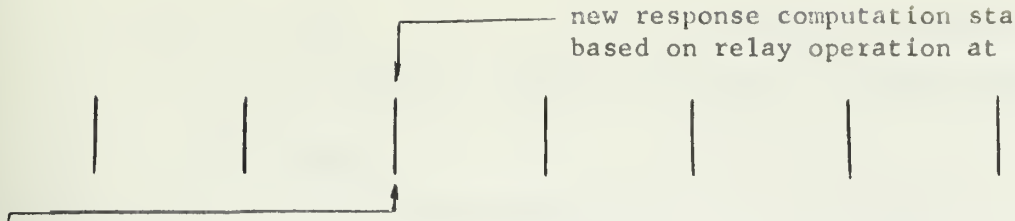


Fig. 7.4. The block diagram of relay servo system.

The variable names used in the program, E(IT), EDELAY, etc., are shown below their mathematical counterparts. Note that the fixed point IT or the floating point T indicates the number of increments, starting with IT = 1:

IT = 1	IT = 2	IT = 3
T = 1.	T = 2.	T = 3.
E(1)	E(2)	E(3)
etc.	etc.	etc.

The method of computation is simply to evaluate the analytic expression for the response at each increment point. The response is always supposed to have started at the time of the last relay operation (TF in increment measure) with initial condition equal to the final values of A(IT) = AF, and DA(IT)/CTACH = DAF at the end of the previous response segment:



if computation at this point shows change of sign of EDELAY from previous value, the relay operates here, TF, AF, and DAF are recorded as the current value of IT, A(IT) and DA(IT).

Analytic response.

The plant transfer function is

$$\frac{A(s)}{B(s)} = \frac{1}{s(s+\alpha)} \tag{7.1}$$

$$B(s) = s^2 A(s) + \alpha s A(s) \tag{7.2}$$

and the differential equation describing the plant is

$$b(t) = \frac{d^2 a(t)}{dt^2} + \alpha \frac{da(t)}{dt} \tag{7.3}$$

with initial conditions a(0) and $\dot{a}(0)$ (AF and DAF in program).

The input is a step V

$$\frac{V}{s} = s^2 A(s) - s a(0) - \dot{a}(0) + \alpha s A(s) - \alpha a(0)$$

$$\therefore A(s) = \frac{V + s^2 a(0) + s\{\dot{a}(0) + \alpha a(0)\}}{s^2(s+\alpha)} \tag{7.4}$$

This can be expressed into partial fractions

$$A(s) = \frac{X}{s^2} + \frac{Y}{s} + \frac{Z}{s+\alpha} = \frac{X(s+\alpha) + Ys(s+\alpha) + Zs^2}{s^2(s+\alpha)}$$

$$= \frac{(Y+Z)s^2 + (X+Y\alpha)s + X\alpha}{s^2(s+\alpha)} \tag{7.5}$$

equating the coefficient of s

$$s^2: \quad Y + Z = a(0) \tag{7.6}$$

$$s^1: \quad X + Y\alpha = \dot{a}(0) + \alpha a(0) \tag{7.7}$$

$$s^0: \quad X\alpha = V \tag{7.8}$$

$$X = \frac{V}{\alpha}$$ = V/ALPHA in program

$$Y = \frac{1}{\alpha}\{\dot{a}(0) + \alpha a(0) - X\}$$ = (DAF + ALPHA-X)/ALPHA

$$Z = a(0) - Y$$ = AF - Y

from equation 7.5, the time solution is

$$a(t) = Xt + Y + Ze^{-\alpha t}$$

= X*TAU + Y + Z*EXPF(-ALPHA*TAU) in program

$$\dot{a}(t) = X - Z\alpha e^{-\alpha t}$$

= X - Z*ALPHA*EXPF(-ALPAH*TAU) in program

t ⟷ TAU is time from last relay operation

Program accuracy.

1). The time delays can be introduced only as an integral number of increments (each of magnitude DELTAT).

2). The only actual inaccuracy is the fact that the relay can only operate at integral values of IT*. The accuracy, for a given increment size, DELTAT could be fairly easily improved by introducing an interpolation procedure between computation points. The same technique could be used to introduce non-integral number of the delays.

The plant dynamics can be easily changed by re-working the partial fraction expression and the expression for a(t) and $\dot{a}$(t).

The program for the relay servo system is shown below. And one case of its output (t_d = .8 second, k_t = 2.0) is attached to this program.

*This error may be regarded as a variability in the forward path time delay, so that the actual delay may lie between the desired ID1 and ID1+1.

```
..JOB CHIANG      MAX. TIME  3 MIN.
      PROGRAM TLRSOI
C RESPONSE CALCULATIONS FOR RELAY SERVO  SYSTEM WITH DELAYS ID1,ID2 AND
C ID3 IN THE FORWARD,FEEDBACK AND TACHOMETER PATHS, RESPECTIVELY.
C SOLUTION IS OBTAINED FOR A TOTAL TIME OF TMAX SECONDS IN INCREMENTS
C OF DELTAT SECONDS. IN THE PROGRAM BOTH T AND IT ARE THE NUMBER OF
C TIME INCREMENTS PLUS ONE. THAT IS, THE INITIAL TIME IS T=1.. THE THREE
C DELAYS ARE MEASURED BY THE NUMBER OF TIME INCREMENTS. THE SYSTEM
C FORCING FUNCTION IS SPECIFIED BY PROGRAM STATEMENT NUMBER 1, AND TWO
C FOLLOWING CARDS.
C INPUT DATA.
C      IN             ON CARD  1. RIGHT JUSTIFIED IN COLUMNS 1 AND 2
C      ID1,ID2,ID3    ON CARDS 2 THRU 4. RIGHT JUSTIFIED IN COLUMNS 1 THRU 3
C      V              ON CARD  5. IN COLUMNS 1 THRU 10 WITH DECIMAL POINT
C      ALPHA          ON CARD  6. IN COLUMNS 1 THRU 10 WITH DECIMAL POINT
C      CTACH          ON CARD  7. IN COLUMNS 1 THRU 10 WITH DECIMAL POINT
C      TMAX           ON CARD  8. IN COLUMNS 1 THRU 10 WITH DECIMAL POINT
C      DELTAT         ON CARD  9. IN COLUMNS 1 THRU 10 WITH DECIMAL POINT
C      INCPRNT        ON CARD 10. RIGHT JUSTIFIED IN COLUMNS 1 AND 2.
C INCPRNT IS THE NUMBER OF INCREMENTS BETWEEN PRINT OUTS
C V IS THE RELAY VOLTAGE
C ALPHA IS THE PLANT TIME CONSTANT
C CTACH IS THE TACHOMETER CONSTANT
C THE MAXIMUM NUMBER OF INCREMENTS IS 1000. THE NUMBER OF INCREMENTS
C MAY NOT EXCEED 900 IF A GRAPH IS REQUIRED.
C IN IS THE NUMBER OF CASES TO BE SOLVED. NOTE THAT THERE MUST BE ONE
C DATA DECK OF NINE CARDS FOR EACH CASE. THE FIRST CARD CONTAINING IN
C IS NOT REPEATED.PUT IN=1 IF ONLY ONE CASE IS TO BE SOLVED.
C EACH SET OF DATA CARDS MUST CONTAIN THE REQUIRED GRAPH FORMAT CARDS.
C REMOVE CALL GRAPH STATEMENT AND GRAPH FORMAT CARDS IF NO GRAPH IS
C REQUIRED.
      DIMENSION A(1001),B(1001),E(1001),DA(1001),DE(1001),ET(1001),
     1          TIMES(1001),FIN(1001),BS(1001)
      READ 150, IN
  150 FORMAT (I2)
      DO 34 K=1,IN
      READ 100,ID1,ID2,ID3,V,ALPHA,CTACH,TMAX,DELTAT,INCPRNT
      PRINT 200
      PRINT 201
      PRINT 202,ID1,ID2,ID3,V,ALPHA,CTACH,TMAX,DELTAT
  100 FORMAT (I3/I3/I3/F10.1/F10.1/F10.1/F10.1/F10.1/I2)
  200 FORMAT (//)
  201 FORMAT(31H RELAY SERVO SYSTEM INPUT DATA  //)
  202 FORMAT ( 8H ID1    =I3/ 8H ID2    =I3/ 8H ID3    =I3/8H V      =F7.3/
     1         8H ALPHA =F7.3/ 8H CTACH =F7.3/ 8H TMAX  =F7.3,5H SEC./
     2         8H DELTAT=F7.3,5H SEC.)
      PRINT 203
      PRINT 204
  204 FORMAT (6X42HCALCULATED RESPONSE OF RELAY SERVO SYSTEM )
  203 FORMAT (1H1)
      T=1.
      TF=1.
      A(1)=0.
      DA(1)=0.
      AF=0.
      DAF=0.
      EDELAY=0.
      INC=TMAX/DELTAT
      NUM=INC-2
    1 F=1.
      DF=0.
  205 FORMAT (6X30HFORCING FUNCTION = A UNIT STEP)
      FIN(IT)=F
      IF(T-1.) 7,2,7
    2 PRINT 205
      PRINT 200
      PRINT 206
      PRINT 207
  206 FORMAT (7X4HTIME5X7HFORC ING5X6HSYSTEM5X5HPLANT4X5HPLANT7X4HRATE
     1        7X4HRATE)
```

```
  207 FORMAT (16X8HFUNCTION4X5HERROR6X5HINPUT4X6HOUTPUT5X6HOUTPUT
     1          6X5HERROR/)
      E(1)=F
      ET(1)=F
      IF(E(1))3,5,3
    3 IF(ID1)5,4,5
    4 B(1)=SIGNF(V,E(1))
      GO TO 6
    5 B(1)=0.
    6 TIME=0.
      DE(1)=DF
      PRINT 208, TIME,F,ET(1),B(1),A(1),DA(1),DE(1)
  208 FORMAT (6XF6.3,3XF8.4,4XF8.4,3XF6.2,3XF8.4,3XF8.4,3XF8.4)
      DO 33 IT=2,INC
      T=IT
      TIME=(T-1.)*DELTAT
      TIMES(IT)=TIME
      EDELAY1=EDELAY
      J=1
      GO TO 1
    7 TAU=(T-TF)*DELTAT
C TAU IS TIME FROM LAST RELAY OPERATION (OR FROM TIME ZERO)
C TF IS THE TIME OF THE LAST RELAY OPERATION (INCREMENT MEASUREMENT)
      X=B(IT-1)/ALPHA
   50 Y=(DAF+ALPHA*AF-X)/ALPHA
      Z=AF-Y
C AF AND DAF ARE THE PLANT OUTPUT AND RATE OUTPUT AT TIME OF LAST RELAY
C OPERATION. X,Y AND Z ARE THE COEFFICIENTS IN THE PARTIAL FRACTION
C EXPANSION.
      A(IT)=X*TAU+Y+Z*EXPF(-ALPHA*TAU)
      DA(IT)= X-Z*ALPHA*EXPF(-ALPHA*TAU)
      IF(ID2)9,8,9
    8 ADELAY=A(IT)
      GO TO 12
    9 IF(ID2-IT+1)10,10,11
   10 ADELAY=A(IT-ID2)
      GO TO 12
   11 ADELAY=0.
   12 IF(ID3)14,13,14
   13 DADELAY=DA(IT)
      GO TO 17
   14 IF(ID3-IT+1)15,15,16
   15 DADELAY=DA(IT-ID3)
      GO TO 17
   16 DADELAY=0.
   17 E(IT)=F-ADELAY-DADELAY*CTACH
      GO TO (51,31),J
   51 IF(ID1)19,18,19
   18 EDELAY=E(IT)
      GO TO 22
   19 IF(ID1-IT+1)20,20,21
   20 EDELAY=E(IT-ID1)
      GO TO 22
   21 EDELAY=0.
   22 IF(EDELAY)24,23,24
   23 B(IT)=0.
      GO TO 25
   24 B(IT)=SIGNF(V,EDELAY)
   25 IF(XMODF((IT-1),40*INCPRNT))27,26,27
   26 PRINT 203
   27 IF(XMODF((IT-1),10*INCPRNT))29,28,29
   28 PRINT 210
  210 FORMAT (/)
   29 IF(B(IT)-B(IT-1))30,31,30
   30 IF(ID1-IT+1)332,331,332
  331 S=0.
      GO TO 333
  332 S=ABSF(EDELAY)/(ABSF(EDELAY)+ABSF(EDELAY1))
  333 TF=T-S
      AF=A(IT)-S*(A(IT)-A(IT-1))
```

```
      DAF=DA(IT)-S*(DA(IT)-DA(IT-1))
C THE LAST SEVEN STATEMENTS SET THE INITIAL VALUES FOR THE NEXT SEGMENT
C OF THE CALCULATION IF RELAY OPERATION HAS BEEN DETECTED BY STATEMENT
C NUMBER 29. THE TIME OF SWITCHING HAS BEEN FOUND BY LINEAR
C INTERPOLATION AS HAVE AF AND DAF. F SHOULD BE CONTINUOUS EXCEPT AT
C TIME ZERO.
      TIME=(T-S-1.)*DELTAT
      PRINT 209,B(IT),TIME
  209 FORMAT (22H THE RELAY SWITCHED TO F6.1,10H AT TIME = F7.3,5H SEC.)
C WE HAVE NOW TO CORRECT A(IT),B(IT) AND E(IT) FOR SWITCHING AT TF.
      TIME=(T-1.)*DELTAT
      J=2
      TAU=(T-TF)*DELTAT
      X=B(IT)/ALPHA
      GO TO 50
   31 BS(IT)=B(IT)*0.125/ABSF(V)+0.125
      ET(IT)=F-A(IT)
      DE(IT)=DF-DA(IT)
      IF(XMODF((IT-1),INCPRNT))33,320,33
  320 PRINT 208,TIME,F,ET(IT),B(IT),A(IT),DA(IT),DE(IT)
   33 CONTINUE
      CALL GRAPH (NUM,ET,DE,8)
      CALL GRAPH (NUM,A ,DA,8)
      CALL GRAPH (NUM,TIMES,A,8)
      CALL GRAPH (NUM,TIMES,DA,8)
      CALL GRAPH (NUM,TIMES,BS,8)
      CALL GRAPH (NUM, TIMES,ET,8)
   34 PRINT 203
      END
      END

 4
  0
  0
  0
.5
1.
.02
30.
0.05
10
0         +0.00E+00 +0.00E+00            00   00   08   08   1    1    02
02 EDOT VS. E
02 CHIANG
0    0
0         +0.00E+00 +0.00E+00            00   00   08   08   1    1    02
02 CDOT VS.C
02 CHIANG
0    0
0         +4.00E+00 +5.00E-01            04   00   08   08   1    1    02
02 C AND CDOT VS.TIME
02 CHIANG
1    0 C
2    0 CDOT
2    0
3    0 E
  2
  0
  0
.5
1.
.02
30.
0.05
10
0         +0.00E+00 +0.00E+00            00   00   08   08   1    1    02
02 EDOT VS. E
02 CHIANG
0    0
0         +0.00E+00 +0.00E+00            00   00   08   08   1    1    02
```

```
02 CDOT VS. C
02 CHIANG
0     0
0          +4.00E+00 +5.00E-01                04    00    08    08    1     1     02
02 C AND CDOT VS. TIME
02 CHIANG
1     0 C
2     0 CDOT
2     0
3     0 E
 16
  0
  0
.5
1.
2.0
30.
0.05
10
0          +0.00E+00 +0.00E+00                00    00    08    08    1     1     02
02 EDOT VS. E
02 CHIANG
0     0
0          +0.00E+00 +0.00E+00                00    00    08    08    1     1     02
02 CDOT VS. C
02 CHIANG
0     0
0          +4.00E+00 +5.00E-01                04    00    08    08    1     1     02
02 C AND CDOT VS. TIME
02 CHIANG
1     0 C
2     0 CDOT
2     0
3     0 E
 20
  0
  0
.5
1.
2.4
30.
0.05
10
0          +0.00E+00 +0.00E+00                00    00    08    08    1     1     02
02 EDOT VS. E
02 CHIANG
0     0
0          +0.00E+00 +0.00E+00                00    00    08    08    1     1     02
02 CDOT VS.C
02 CHIANG
0     0
0          +4.00E+00 +5.00E-01                04    00    08    08    1     1     02
02 C AND CDOT VS.TIME
02 CHIANG
1     0 C
2     0 CDOT
2     0
3     0 E

..END
```

RELAY SERVO SYSTEM INPUT DATA

```
ID1    = 16
ID2    =  0
ID3    =  0
V      =    .500
ALPHA  =   1.000
CTACH  =   2.000
TMAX   = 30.000 SEC.
DELTAT =    .050 SEC.
```

CALCULATED RESPONSE OF RELAY SERVO SYSTEM
FORCING FUNCTION = A UNIT STEP

TIME	FORCING FUNCTION	SYSTEM ERROR	PLANT INPUT	PLANT OUTPUT	RATE OUTPUT	RATE ERROR
.000	1.0000	1.0000	.00	.0000	.0000	.0000
.500	1.0000	1.0000	.00	.0000	.0000	.0000
THE RELAY SWITCHED TO .5 AT TIME = .800 SEC.						
1.000	1.0000	.9906	.50	.0094	.0906	-.0906
1.500	1.0000	.9017	.50	.0983	.2517	-.2517
2.000	1.0000	.7494	.50	.2506	.3494	-.3494
2.500	1.0000	.5587	.50	.4413	.4087	-.4087
THE RELAY SWITCHED TO -.5 AT TIME = 2.879 SEC.						
3.000	1.0000	.3517	-.50	.6483	.3302	-.3302
3.500	1.0000	.2750	-.50	.7250	.0035	-.0035
4.000	1.0000	.3269	-.50	.6731	-.1946	.1946
THE RELAY SWITCHED TO .5 AT TIME = 4.050 SEC.						
4.500	1.0000	.3691	.50	.6309	.0476	-.0476
5.000	1.0000	.2971	.50	.7029	.2256	-.2256
5.500	1.0000	.1551	.50	.8449	.3336	-.3336
THE RELAY SWITCHED TO -.5 AT TIME = 5.608 SEC.						
6.000	1.0000	.0382	-.50	.9618	.0748	-.0748
6.500	1.0000	.0621	-.50	.9379	-.1514	.1514
THE RELAY SWITCHED TO .5 AT TIME = 6.907 SEC.						
7.000	1.0000	.1707	.50	.8293	-.1994	.1994
7.500	1.0000	.1959	.50	.8041	.0758	-.0758
8.000	1.0000	.1128	.50	.8872	.2427	-.2427
THE RELAY SWITCHED TO -.5 AT TIME = 8.348 SEC.						
8.500	1.0000	-.0251	-.50	1.0251	.2033	-.2033
9.000	1.0000	-.0518	-.50	1.0518	-.0734	.0734
9.500	1.0000	.0304	-.50	.9696	-.2413	.2413
THE RELAY SWITCHED TO .5 AT TIME = 9.700 SEC.						
10.000	1.0000	.1377	.50	.8623	-.0838	.0838
10.500	1.0000	.1174	.50	.8826	.1459	-.1459
11.000	1.0000	.0067	.50	.9933	.2853	-.2853
THE RELAY SWITCHED TO -.5 AT TIME = 11.104 SEC.						
11.500	1.0000	-.0899	-.50	1.0899	.0429	-.0429
12.000	1.0000	-.0535	-.50	1.0535	-.1707	.1707
THE RELAY SWITCHED TO .5 AT TIME = 12.476 SEC.						
12.500	1.0000	.0667	.50	.9333	-.2767	.2767
13.000	1.0000	.1223	.50	.8777	.0289	-.0289
13.500	1.0000	.0577	.50	.9423	.2143	-.2143
THE RELAY SWITCHED TO -.5 AT TIME = 13.868 SEC.						
14.000	1.0000	-.0715	-.50	1.0715	.2027	-.2027
14.500	1.0000	-.0980	-.50	1.0980	-.0738	.0738

15.000	1.0000	-.0157	-.50	1.0157	-.2415	.2415
THE RELAY SWITCHED TO .5 AT TIME = 15.247 SEC.						
15.500	1.0000	.1031	.50	.8969	-.1198	.1198
16.000	1.0000	.0970	.50	.9030	.1241	-.1241
16.500	1.0000	-.0051	.50	1.0051	.2720	-.2720
THE RELAY SWITCHED TO -.5 AT TIME = 16.634 SEC.						
17.000	1.0000	-.1059	-.50	1.1059	.0551	-.0551
17.500	1.0000	-.0743	-.50	1.0743	-.1633	.1633
18.000	1.0000	.0433	-.50	.9567	-.2958	.2958
THE RELAY SWITCHED TO .5 AT TIME = 18.016 SEC.						
18.500	1.0000	.1128	.50	.8872	.0074	-.0074
19.000	1.0000	.0566	.50	.9434	.2012	-.2012
THE RELAY SWITCHED TO -.5 AT TIME = 19.401 SEC.						
19.500	1.0000	-.0711	-.50	1.0711	.2247	-.2247
20.000	1.0000	-.1063	-.50	1.1063	-.0604	.0604
20.500	1.0000	-.0292	-.50	1.0292	-.2334	.2334
THE RELAY SWITCHED TO .5 AT TIME = 20.785 SEC.						
21.000	1.0000	.0943	.50	.9057	-.1445	.1445
21.500	1.0000	.0979	.50	.9021	.1091	-.1091
22.000	1.0000	.0017	.50	.9983	.2629	-.2629
THE RELAY SWITCHED TO -.5 AT TIME = 22.169 SEC.						
22.500	1.0000	-.1059	-.50	1.1059	.0744	-.0744
23.000	1.0000	-.0819	-.50	1.0819	-.1516	.1516
23.500	1.0000	.0311	-.50	.9689	-.2887	.2887
THE RELAY SWITCHED TO .5 AT TIME = 23.553 SEC.						
24.000	1.0000	.1113	.50	.8887	-.0112	.0112
24.500	1.0000	.0624	.50	.9376	.1899	-.1899
THE RELAY SWITCHED TO -.5 AT TIME = 24.937 SEC.						
25.000	1.0000	-.0637	-.50	1.0637	.2507	-.2507
25.500	1.0000	-.1090	-.50	1.1090	-.0447	.0447
26.000	1.0000	-.0382	-.50	1.0382	-.2238	.2238
THE RELAY SWITCHED TO .5 AT TIME = 26.321 SEC.						
26.500	1.0000	.0881	.50	.9119	-.1684	.1684
27.000	1.0000	.1011	.50	.8989	.0946	-.0946
27.500	1.0000	.0106	.50	.9894	.2541	-.2541
THE RELAY SWITCHED TO -.5 AT TIME = 27.705 SEC.						
28.000	1.0000	-.1031	-.50	1.1031	.0953	-.0953
28.500	1.0000	-.0873	-.50	1.0873	-.1390	.1390
29.000	1.0000	.0206	-.50	.9794	-.2810	.2810
THE RELAY SWITCHED TO .5 AT TIME = 29.089 SEC.						
29.500	1.0000	.1104	.50	.8896	-.0300	.0300

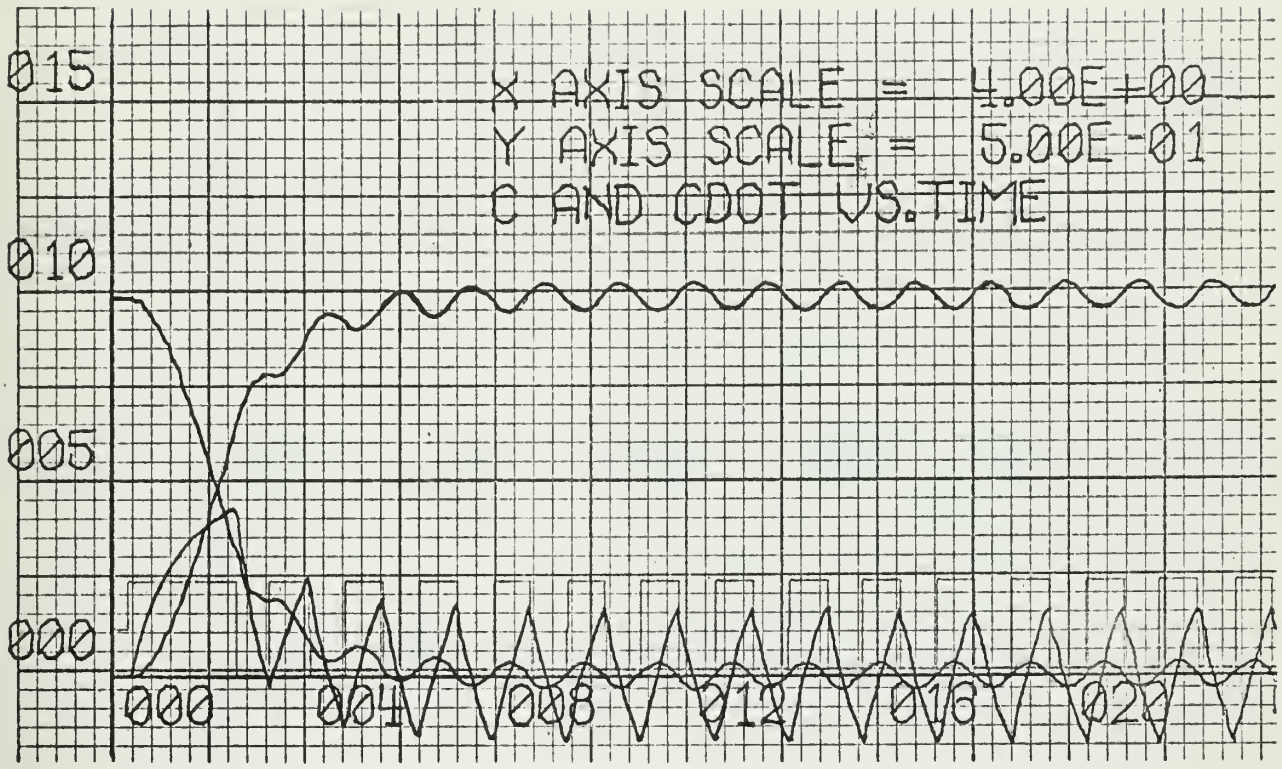
7.3 Comparison of the results obtained from the digital computer and from the analog computer.

The result obtained from the digital computer for the system with time delay .4 second has been checked with analog computer. They are quite in agreement with each other as shown in Fig. 7.5 and 7.6.

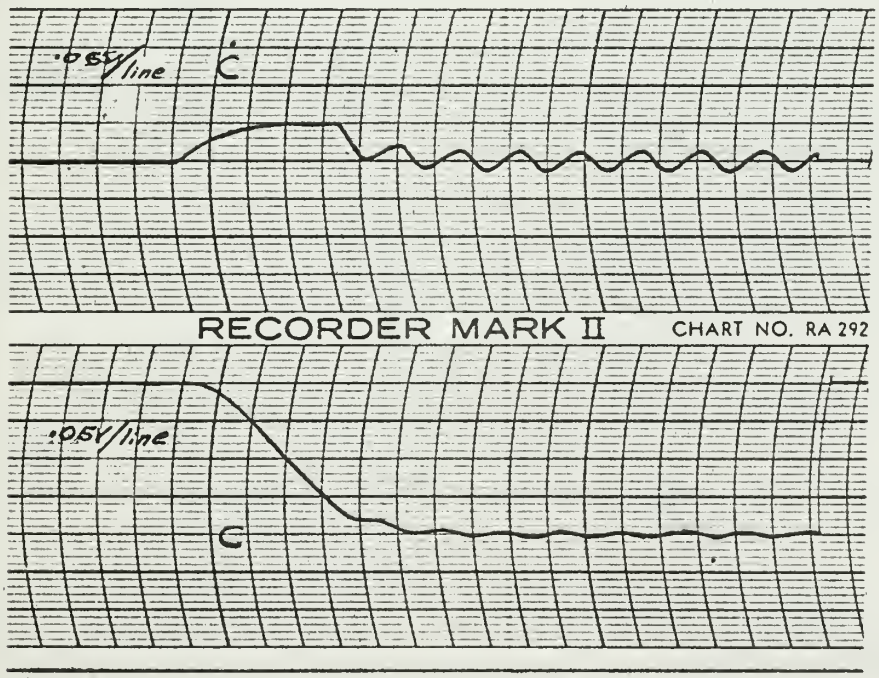
comparison of the results

	by digital computer	by analog computer
overshoot and undershoot	3.5	2.6
rise time	5.9	5.7

Because the digital computer can give an accurate and exact value for any desired instant, therefore, the digital computer is selected for conducting most of the experiment work, and the analog computer is used to check the results obtained by digital computer. They are quite in agreement.



(a) results obtained from digit computer



(b) results obtained from analog computer

Fig. 7-5 Output, velocity of the system, t_d = .4 second

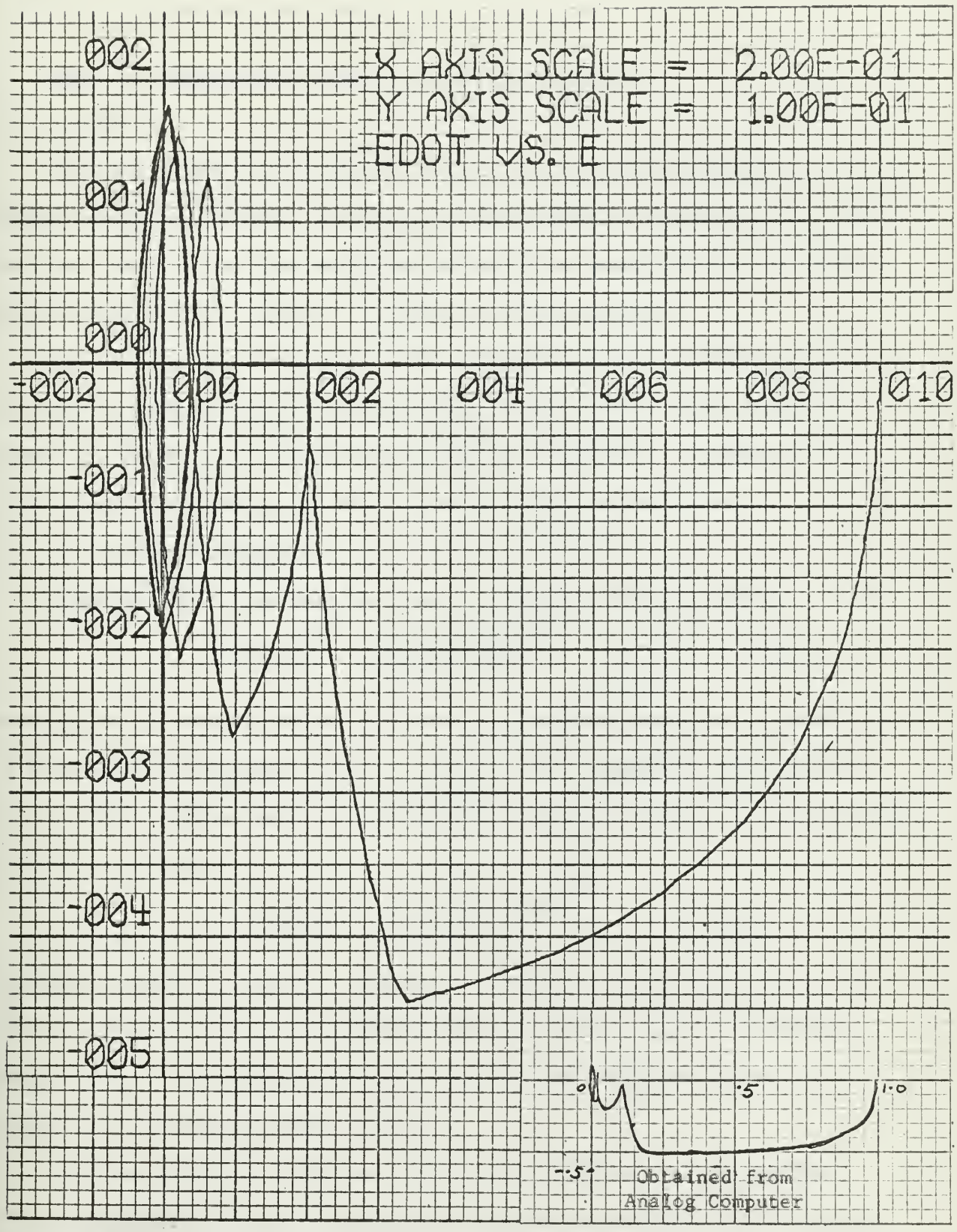


Fig. 7-6 . Phase trajectory of the system with t_d = .4 second

8. Effect of transportation lag on relay servo system.

8.1. General

From the proceeding sections, we know that the time delay has an adverse effect on control systems, and which is unavoidable in nearly all the control systems. Because the time delay may basically be represented by an infinite series of terms, this creates an infinite number of roots.

Later on in this section, it will be noted that the time lag results in successive overshooting and undershooting of the correct final value, representing serious oscillations about the final position. In addition to this it will also change the root locus and phase trajectories.

8.2 The frequency response.

From Fig. 4.6(c) we have the transfer function with G_D removed

$$G(S) = \frac{(1 + K_t S)e^{-ST}}{S(S+1)} \qquad (8.1)$$

where K_t and T are variables. In case of time delay equal to .5 second, it is found that the tachometer gain, $k_t = 1.5$, will compensate the system to an acceptable optimum operation, this will be seen in section 9, now equation 8.1 can be written as

$$G(s) = \frac{(1 + 1.5\,S)e^{-.5S}}{S(S+1)} \qquad (8.2)$$

The Bode diagram for the above transfer function is shown in Fig. 8.1. So far as the linear system is concerned, the gain margin for .5 second time delay system is approximately 1.6 db, but for the undelayed system, it is stable for all gains and never unstable.

The time delay has introduced additional branches into the diagram. These branches appear due to the periodic nature of the time delay, and the increasing phase shift introduced at higher frequencies. It is evident that

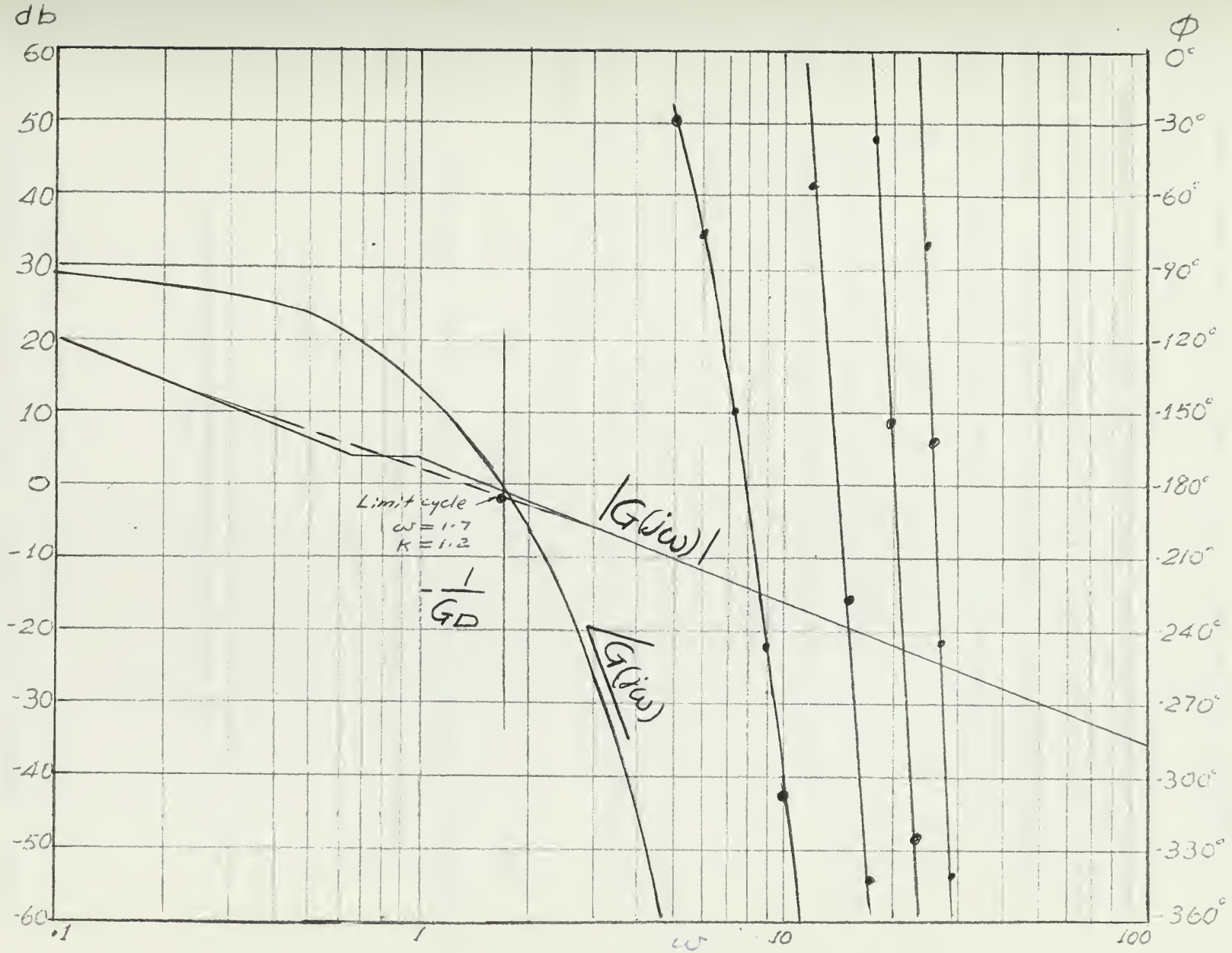


Fig. 8-1 Bode diagram of Eq. 8-2.

the gain margin for each of the branches increases. This is analogous to the statement that the roots of the primary branch would go unstable first if the gain were raised.

The gain vs. $\log \omega$ curve of the delayed system is identical to that of the undelayed system, since the time delay introduced no additional gain to the system. The phase angle vs. $\log \omega$ curve of the delayed system is drastically changed since the equation for the phase angle is

$$\phi = -(90^\circ + \tan^{-1}\omega + \omega T) + \tan^{-1}(1.5\omega) \quad (8.3)$$

The ωT term in the above equation has a predominate effect to the phase angle, it introduces additional phase shift and extra branches.

The limit cycle occurs at $\omega = 1.7$, $k = 1.20$.

8.3 The root locus.

The root locus of any system will be greatly changed when a transportation lag was introduced. As a matter of fact, for an undelayed system the transfer function in Equation(8.2) becomes

$$G(S) = \frac{1.5(S + .67)}{S(S + 1)} \quad (8.4)$$

no limit cycle exists, because the root locus never intersects the imaginary axis as shown in Fig. 8.2, which shows a stable system for all gains.

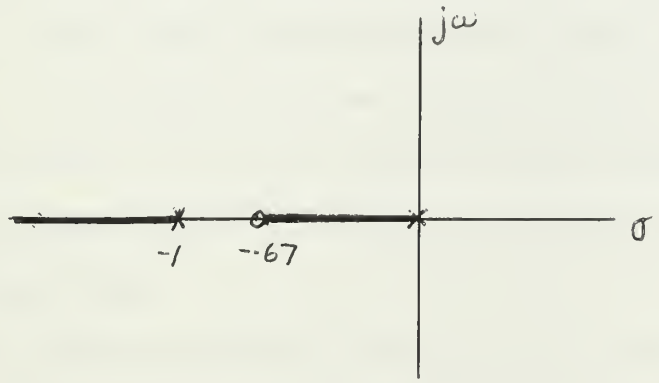


Fig. 8.2. Root locus of Eq. 8.4.

But when time delay is introduced this situation will be changed. The root

locus for Eq. 8.2 (.5 second delayed system) is shown in Fig. 8.3. The construction details for this root locus were described in section 4.4.2.

Fig. 8.3 illustrates the effect of time delay on the root locus. It is apparent that this system will go unstable with only a small increase in gain. The primary branch stability limit before the root locus corsses the imaginary axis are k = 1.2., and ω = 1.76 rad/sec., where the limit cycle occurs, this almost checks with Bode diagram where k = 1.20, ω = 1.7 rad/sec.

Because time delay introduces periodically phase shift of 360 degrees, so there presents additional (upper) branches when at high frequencies. It is noted that for the same gain the additional roots on the upper branches are attenuated more than the roots on the primary branch. This substantiates the Bode diagram findings, which shows that the roots on the primary branch phase shift would cause instability.

8.4 Transient response.

Because of time delay introduced in the system, there is no graphical method that can be easily used to produce the transient response of any system. The transient response of the delayed relay servo system (second order) for various delay time and various k_t (tachometer gain) are shown in Fig. 8.4. In Fig. 8.4(a) there are six curves under the same tachometer gain (k_t = .02); curve (1) is an undelayed system, the rest of the five curves are with .1, .25, .5, .8 and 1.0 second delay time respectively. For 1 second time delay system it presents an overshoot and undershoot of approximately 60 percent.

Table 8.1 is the data obtained for M_{p_t}, t_p, t_r under different values of k_t and t_d. Plotting these data into curves it is found that they appear as a straight line. In Fig. 8.5 are shown M_{p_t} and t_p versus delay time, in

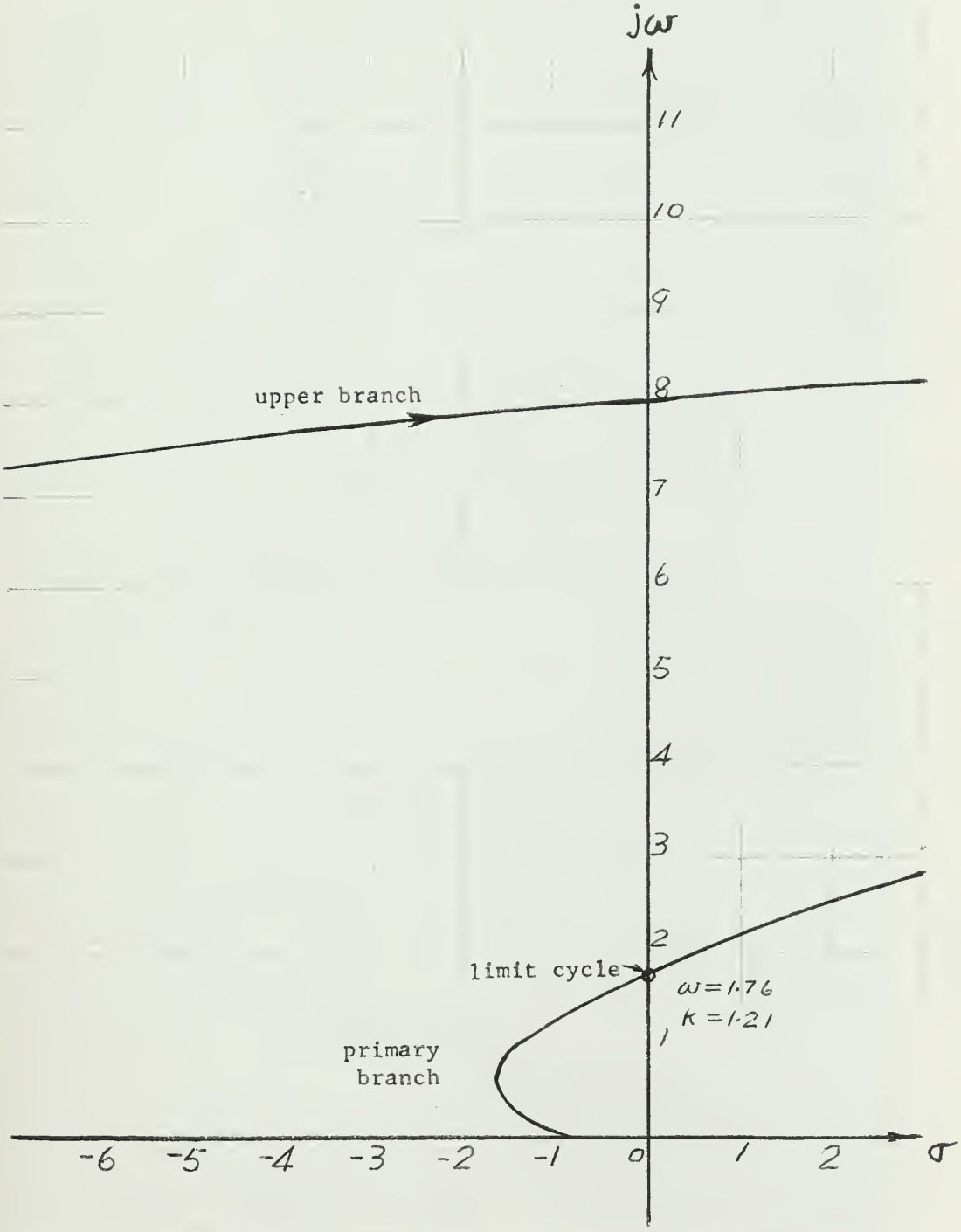


Fig. 8-3 Root locus of relay servo system with time delay of .5 second.

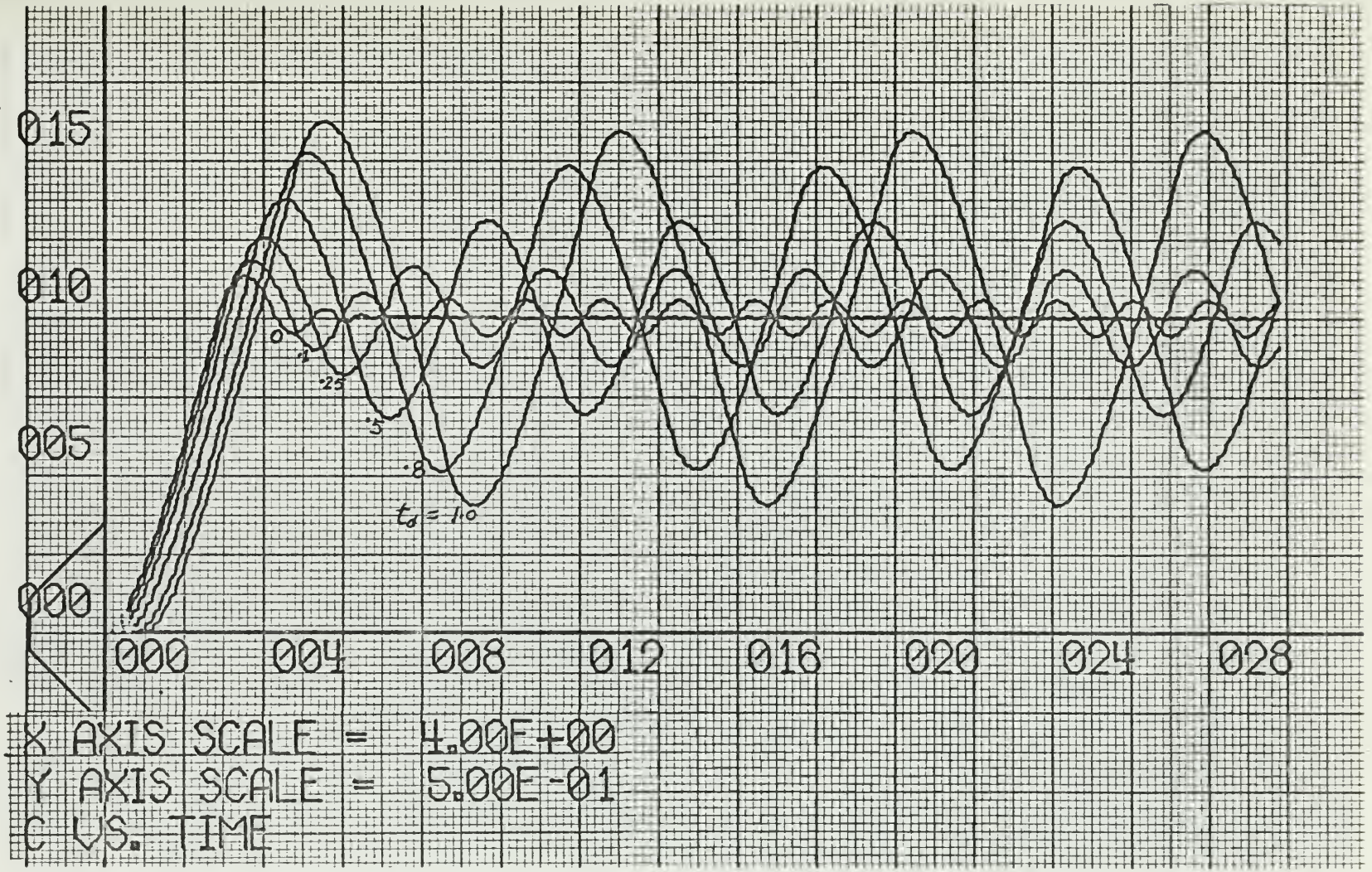


Fig. 8-4a. Response of relay servo system t_d=0, .1, .25, .5, .8, 1.0 second respectively, k_t=.02

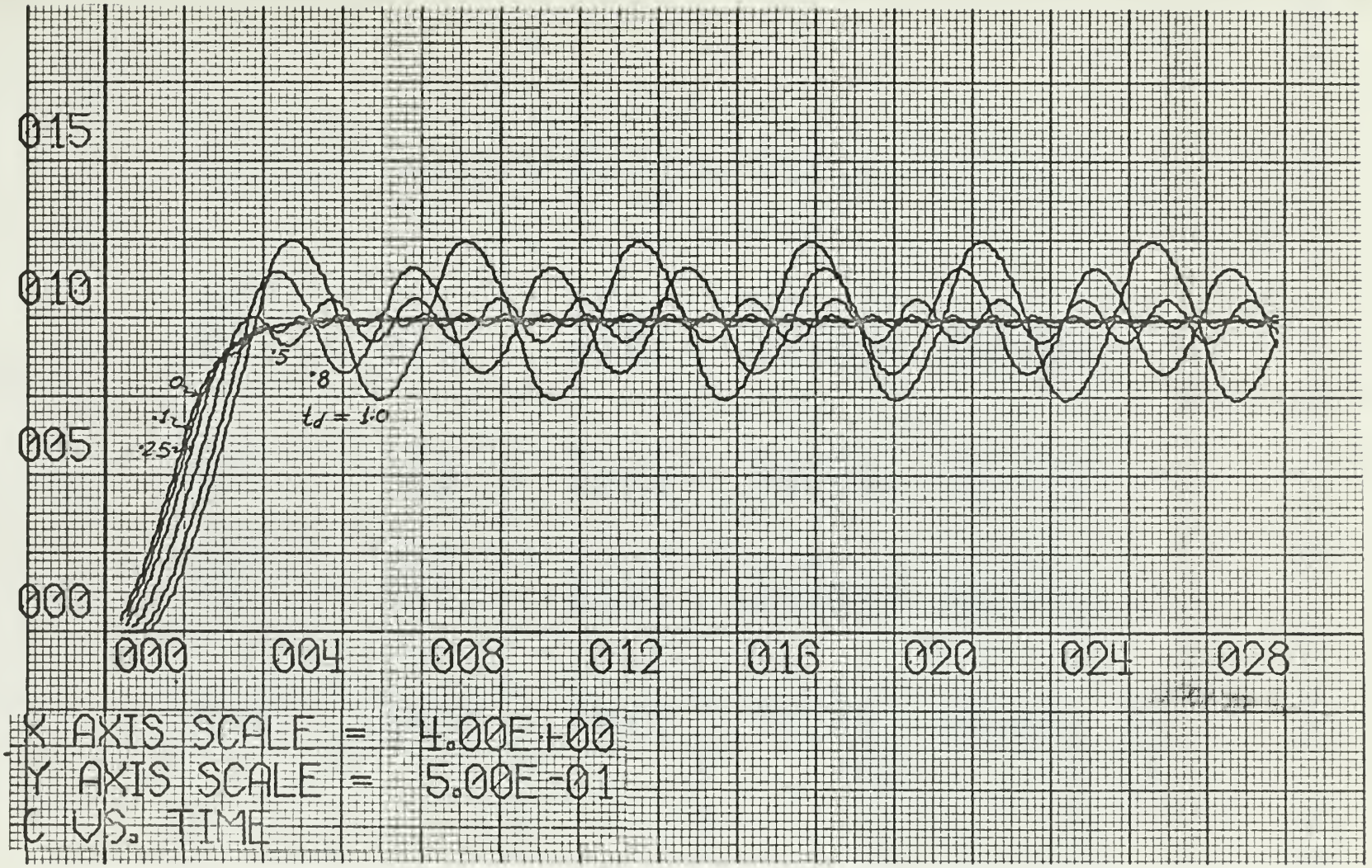


Fig. 8-4b . Response of relay servo system with t_d= 0, .1, .25, .5, .8, 1.0 second respectively, k_t=.8

Table 8.1. Data for first overshoot t_p and M_{p_t} under various time delays for a step input.

K_t	T_d	M_{p_t}	t_p	t_r	amplitude of limit cycle
.02	.0	1.1285	3.50	2.828	no limit cycle
"	.1	1.1870	3.70	2.95	.08
"	.25	1.2510	4.10	3.095	.1625
"	.5	1.3750	4.60	3.350	.2575
"	.8	1.5300	5.10	3.650	.4910
"	1.0	1.6500	5.50	3.850	.5900
.2	.0	1.0370	3.35	2.900	no limit cycle
"	.1	1.0870	3.51	2.950	negligible
"	.25	1.1630	3.81	3.080	.057
"	.5	1.2880	4.30	3.350	.213
"	.8	1.4370	4.95	3.650	.388
"	1.0	1.5310	5.40	3.850	.49
.4	.0	/	/	/	no limit cycle
"	.1	/	/	/	"
"	.25	1.0600	3.70	3.200	.024
"	.5	1.1650	4.20	3.420	.125
"	.8	1.3280	4.80	3.700	.302
"	1.0	1.4300	5.20	3.900	.390
.8	.0	/	/	/	no limit cycle
"	.1	/	/	/	"
"	.25	/	/	/	.015
"	.5	1.0000	3.70	3.60	.061
"	.8	1.1500	4.39	3.65	.153
"	1.0	1.2495	4.71	3.85	.2495
1.0	.0	/	/	/	no limit cycle
"	.1	/	/	/	"
"	.25	/	/	/	.014
"	.5	/	/	/	.055
"	.8	1.0550	4.10	3.70	.137
"	1.0	1.1550	4.60	3.85	.205

_____ t_p (time for first overshoot)

- - - - M_p (magnitude for first overshoot)

curve	k_t
1	.02
2	.2
3	.4
4	.8
5	1.0

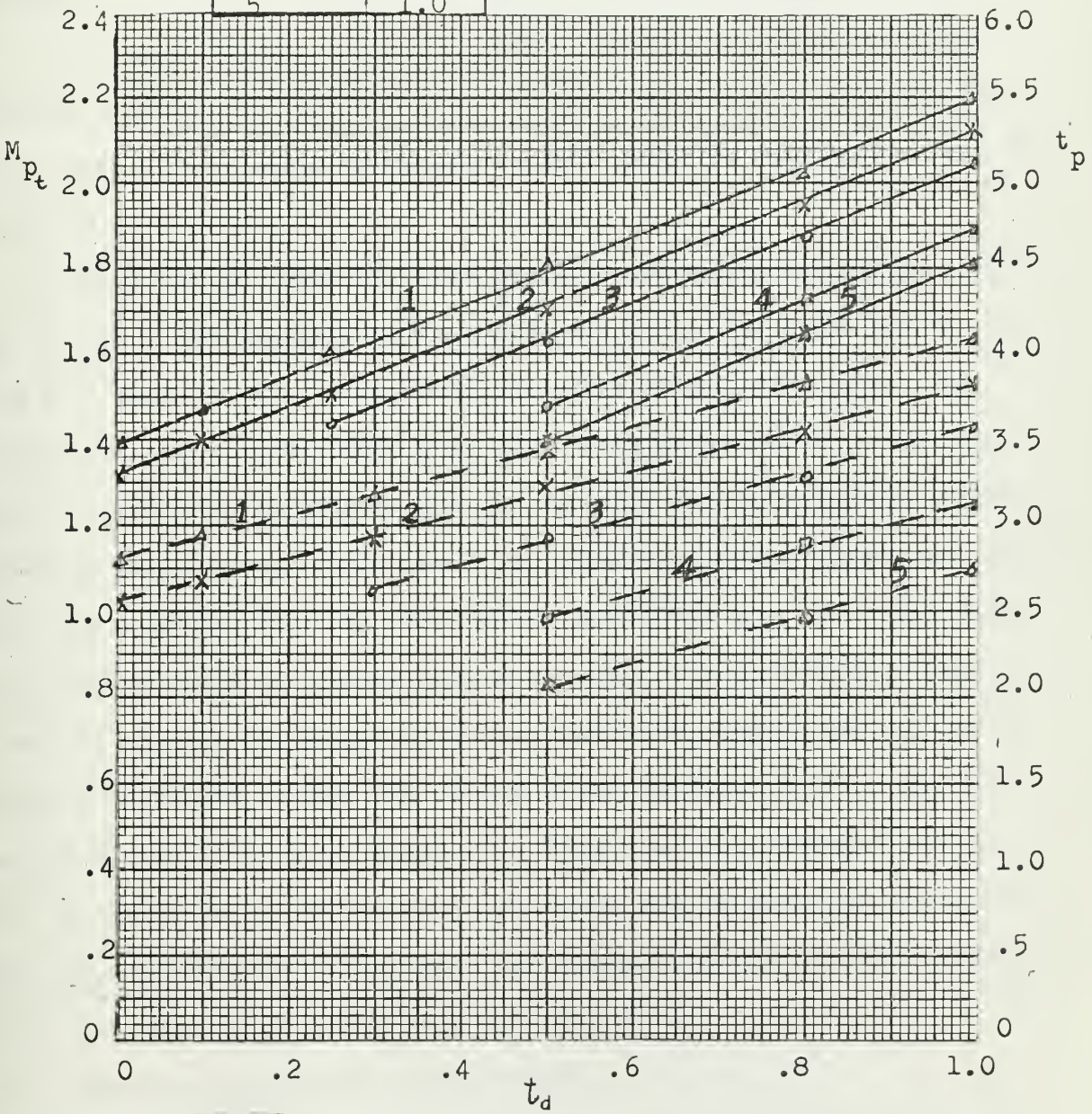


Fig. 8-5 M_{p_t} and t_p vs. delay time

Fig. 8.6 is shown the rise time versus delay time, there it shows that the effect of time delay on the transient response is to increase the overshoot and undershoot and cause continuously oscillations and make the rise time longer, it turns out that the response of the system can never be settled.

It is noted that M_{p_t}, t_p, t_r increase as the delay time increases. But when we increase the value of k_t they decrease. And to a certain value of k_t some of them become acceptable optimum operation this will be seen in section 9.2.

In Fig. 8.7 are shown the plant output, rate output and system error. Fig. 8.7(a) is the response of the undelayed system, the plant output settled at 7 seconds; Fig. 8.7(b) shows the plant output, rate output and system error for a system with delay time of .5 second and $k_t = .02$, the output continuously oscillates with an overshoot and undershoot of approximately 30 percent; Fig. 8.7(c) shows the plant output rate output and system error for a system with time delay of one second and $k_t = .02$. The plant output continuously oscillates with an overshoot and undershoot of approximately 60 percent, the frequency of oscillation is slower than that of .5 second time delayed system; Fig. 8.7(d) shows the plant output, rate output and system error for a system with delay time of 2.5 seconds and $k_t = .02$, the plant output continuously oscillates with an overshoot and undershoot of approximately 135 percent, the frequency of oscillations is much slower than that of .5 second time delayed system.

8.5 Time delay in the phase plane diagram.

The effect of time delay in application or removal of the applied torque after the control has signaled that the operation should be performed. It will be of interest to study the effect in terms of the phase plane diagram.

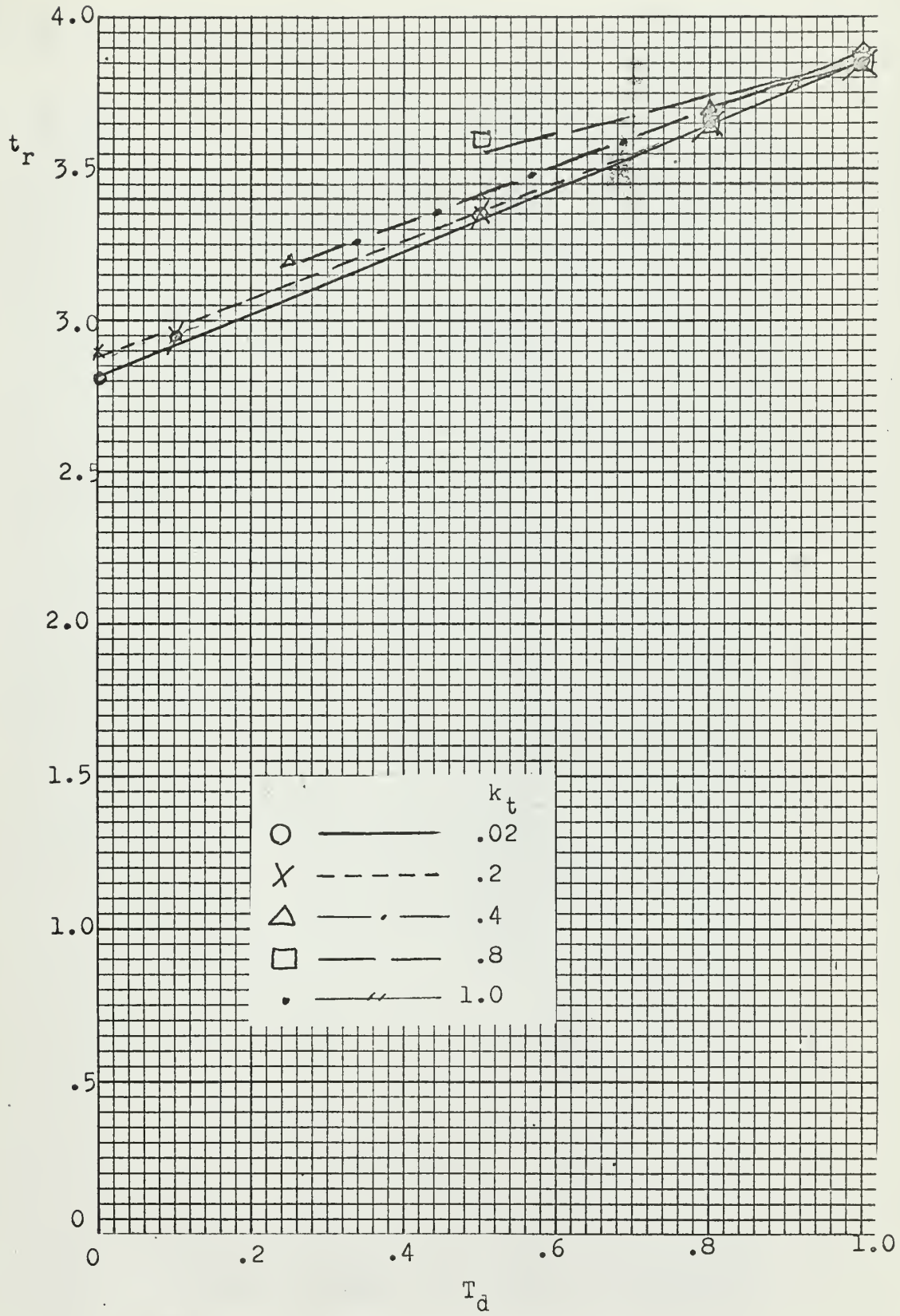


Fig. 8-6 Time rise vs. delay time

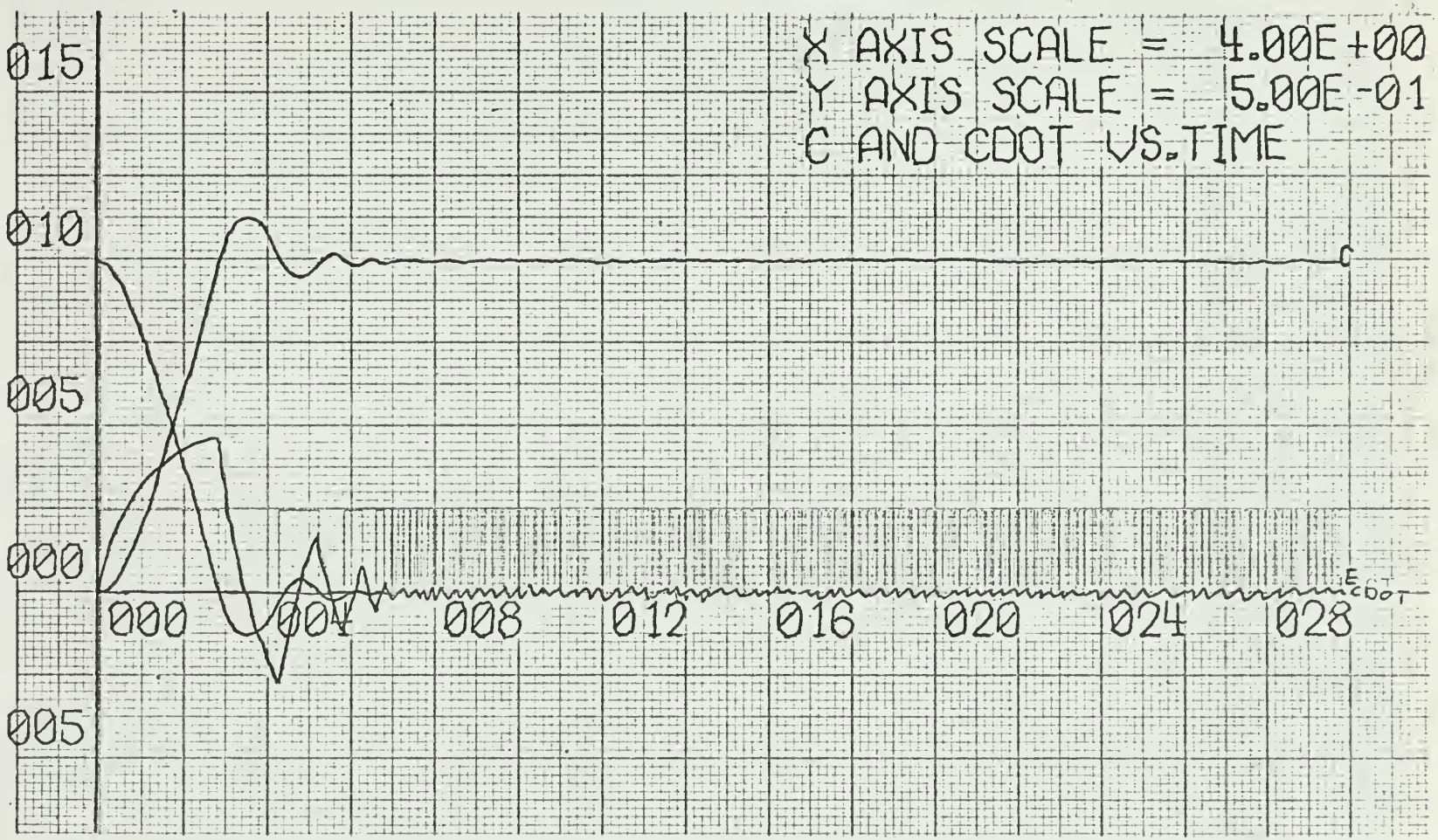


Fig. 8-7a The plant output, velocity and system error of an undelayed relay servo system

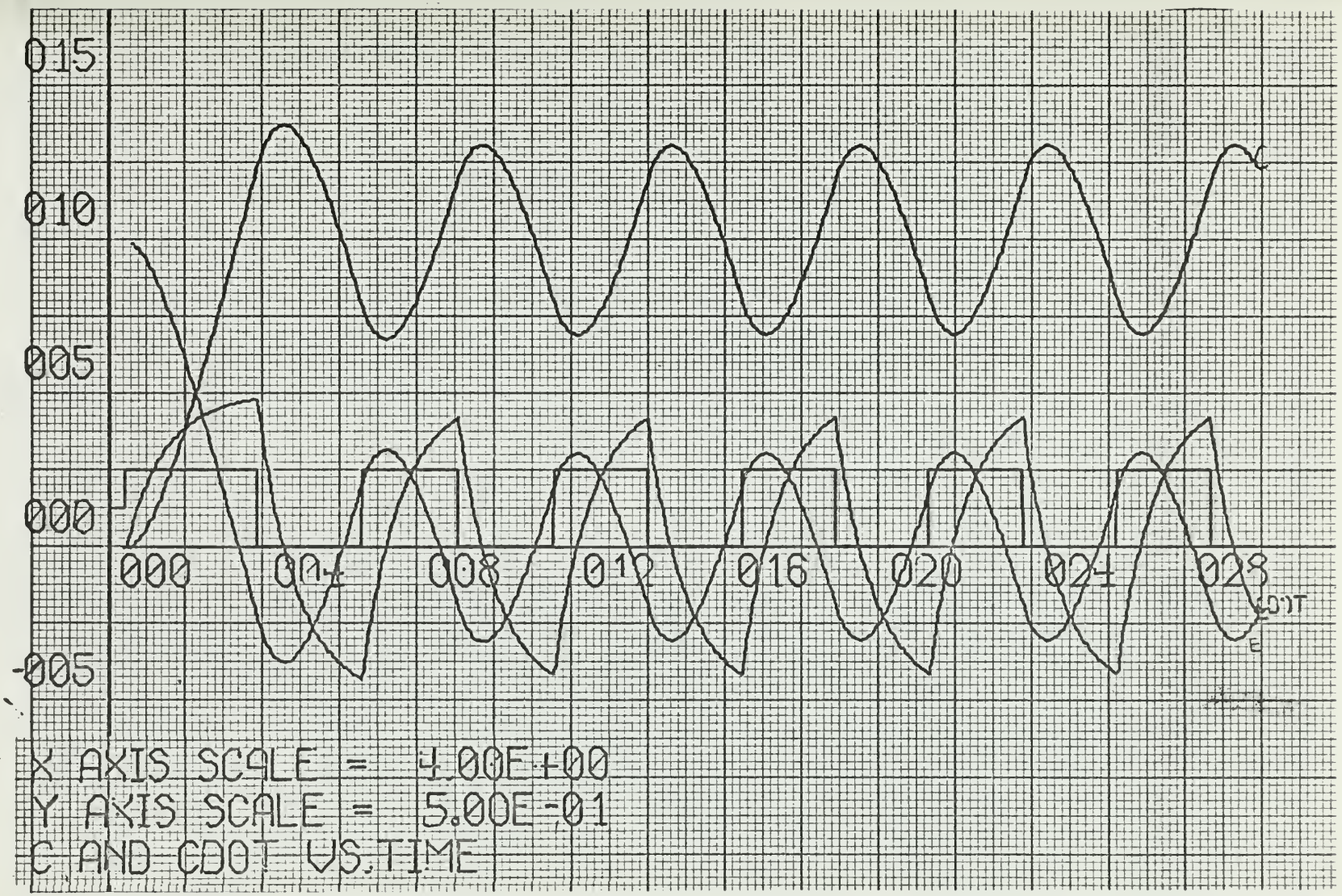


Fig. 8-7b. The plant output, velocity and error of relay servo system with a time delay of .5 second and $k_t = .02$.

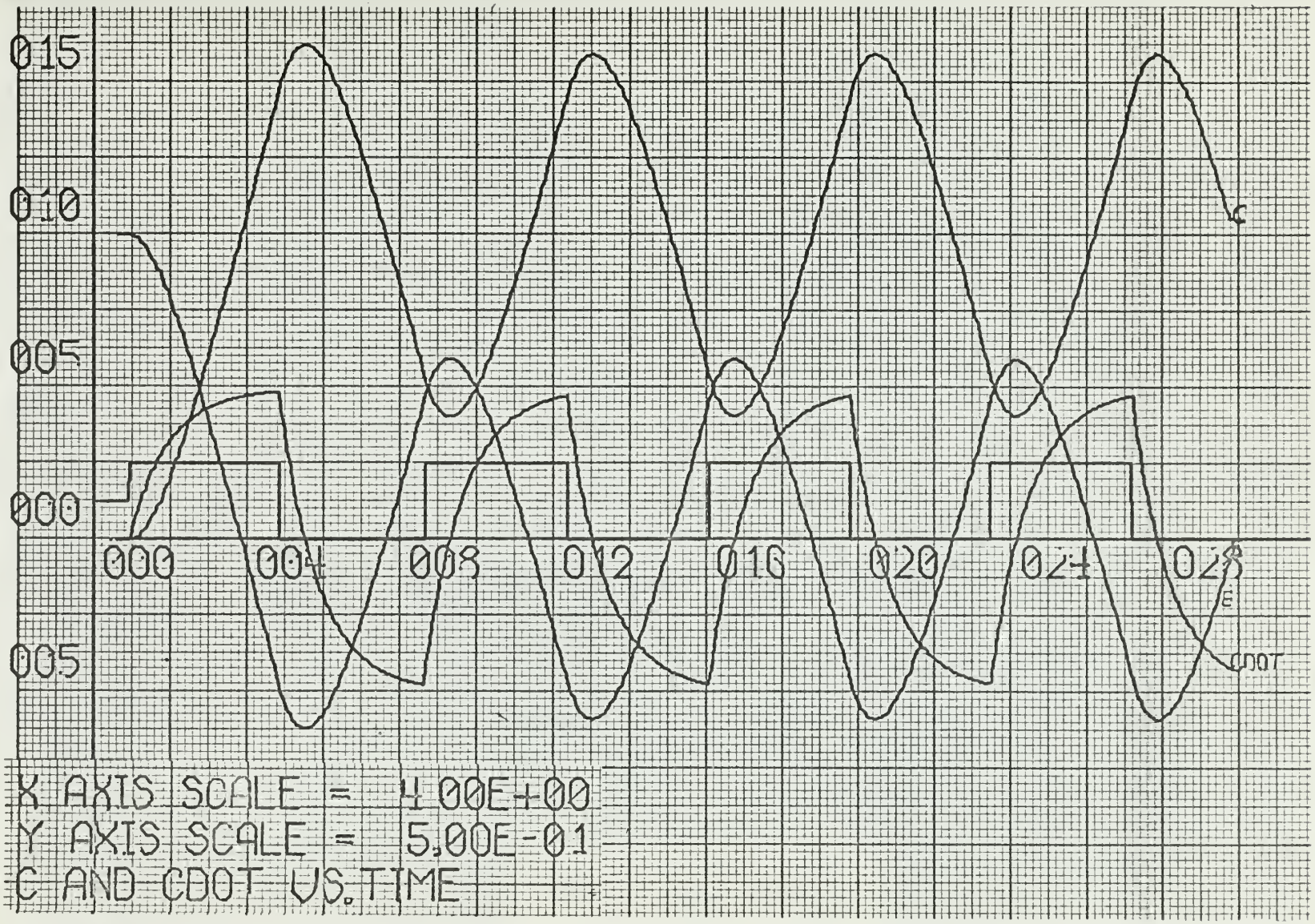


Fig. 8-7c. The plant output, velocity and system error of a relay servo system with a time delay of one second and $k_t = .02$

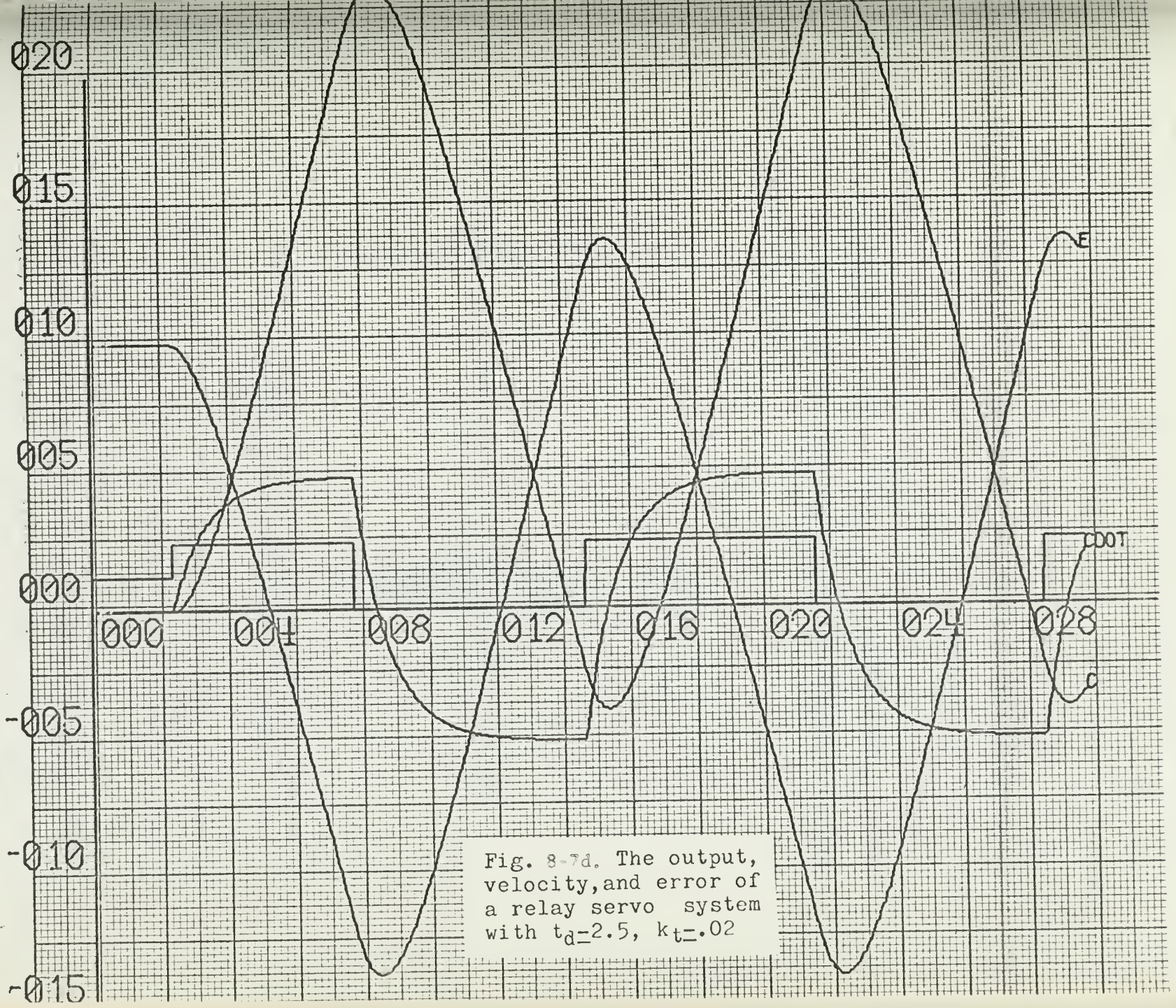


Fig. 8-7d. The output, velocity, and error of a relay servo system with $t_d=2.5$, $k_t=.02$

Because of the time delay the trajectory in the phase space will be prolonged beyond the point $c' = R'$ by an amount such that the difference in coordinates is just equal to the time delay. That is, in Fig. 8.8 the torque should reverse at a, but it does not reverse until B is reached, and $(C'_B - C'_A)$ $(v'_B - v'_A)/2$ is equal to t'_d. Similarly, the torque is reversed at point D instead of at c, and at F instead of at E. It is apparent that a curve could be drawn through points such as B and F so that the switching of torque direction would take place when the trajectory crossed this line, and the line from B to D would play the same role as the line through R' for the control without time delay.

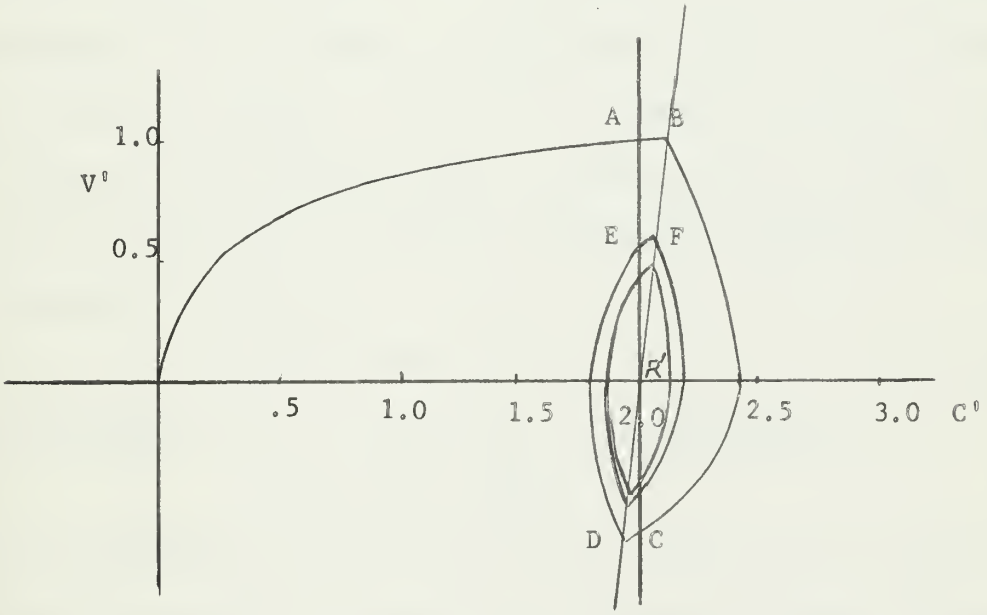


Fig. 8.8. Normalized phase plane diagram for relay servo system with a time delay of .1 second, for a step unit.

The effect of time delay is thus to skew the switching line from the verticle in the phase space, and quite approximately the angle of skew is $-\tan^{-1} t_d$ and is in clockwise direction.

It is seen in Fig. 8.8 that the trajectory appears to repeat itself, since the last two oscillations are nearly identical. The control oscillates

with a constant amplitude on the trajectory. Such a repetive diagram representing the steady state oscillation is called limit cycle. From the experimental work it shows that the relay controls with time delay and no dead zone exhibits the defect of continuous oscillations.

In Fig. 8.9 is shown a family of trajectories so arranged that the initial and final points of each are symmetrically located with respect to the origin. Thus each trajectory represents just such a limit cycle, and since the time delay for each is fixed, the locus of the end points of these limit cycles measures the time delay required to obtain the amplitude of oscillation. As would be expected, the larger the time delay the larger is the magnitude of the oscillation, and with that, the lower is the frequency of oscillation as indicated in Fig. 8.5. In Fig. 8.10 is shown a plot of time delay and oscillation amplitude. It turns out to be a straight line, therefore, we can obtain an equation to express the relation of oscillation amplitude with time delay as

$$A_{osc} = .6\, t_d \qquad (8.4)$$

the above equation is for a second order system specified by

$$G(S) = \frac{1}{S(S+1)}$$

It is interesting to see the relay type control that can continuosuly cycle when a time delay is present. The relay type of control always oscillates about the equilibrium position. For any time delay the system would always oscillate, and with a magnitude of oscillation depending on the time delay. The relay type control can never be unstable in the sense that the amplitude of oscillation will continue to increase without limit. It is obvious that this is the advantage of the relay servo system.

In Fig. 8.11(a) the phase plane of the relay servo system with a time delay of zero and .5 seconds is plotted on the same sheet. In Fig. 8.11(b)

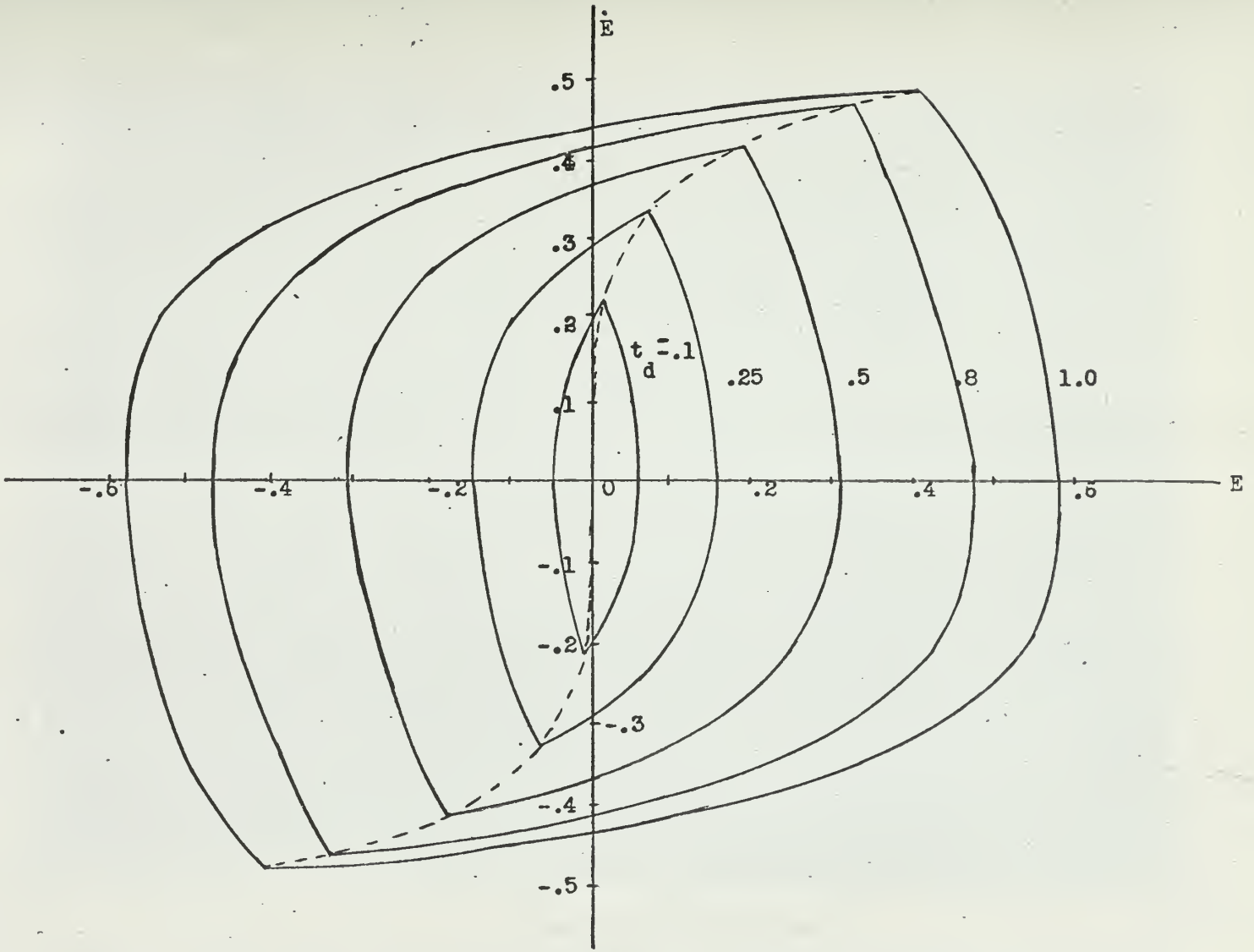


Fig. 8-9. Phase plane trajectory showing loci of limit cycle for oscillation of constant amplitude with various time delay

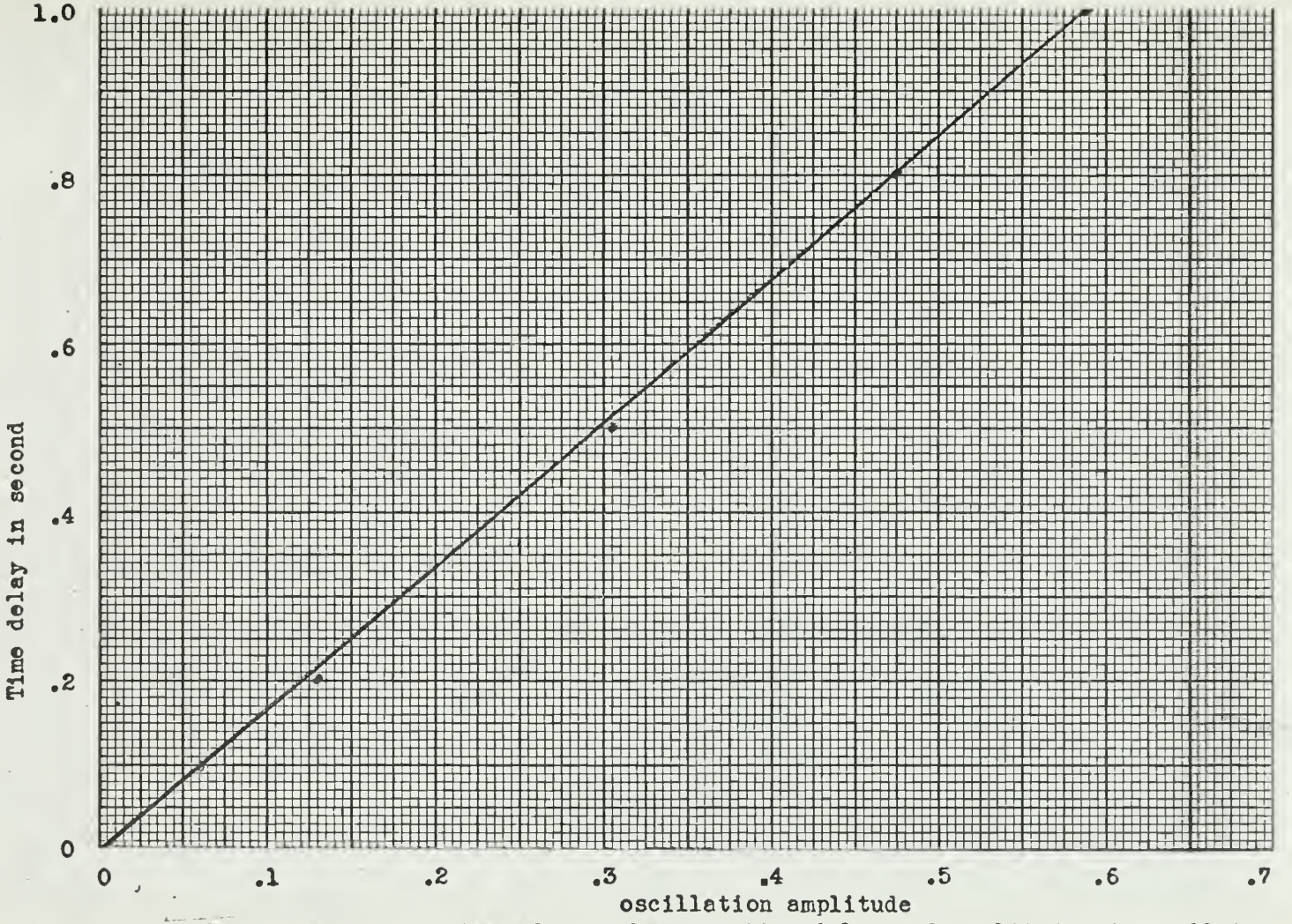


Fig. 8-10. Curve showing the relation between time delay and amplitude of oscillation

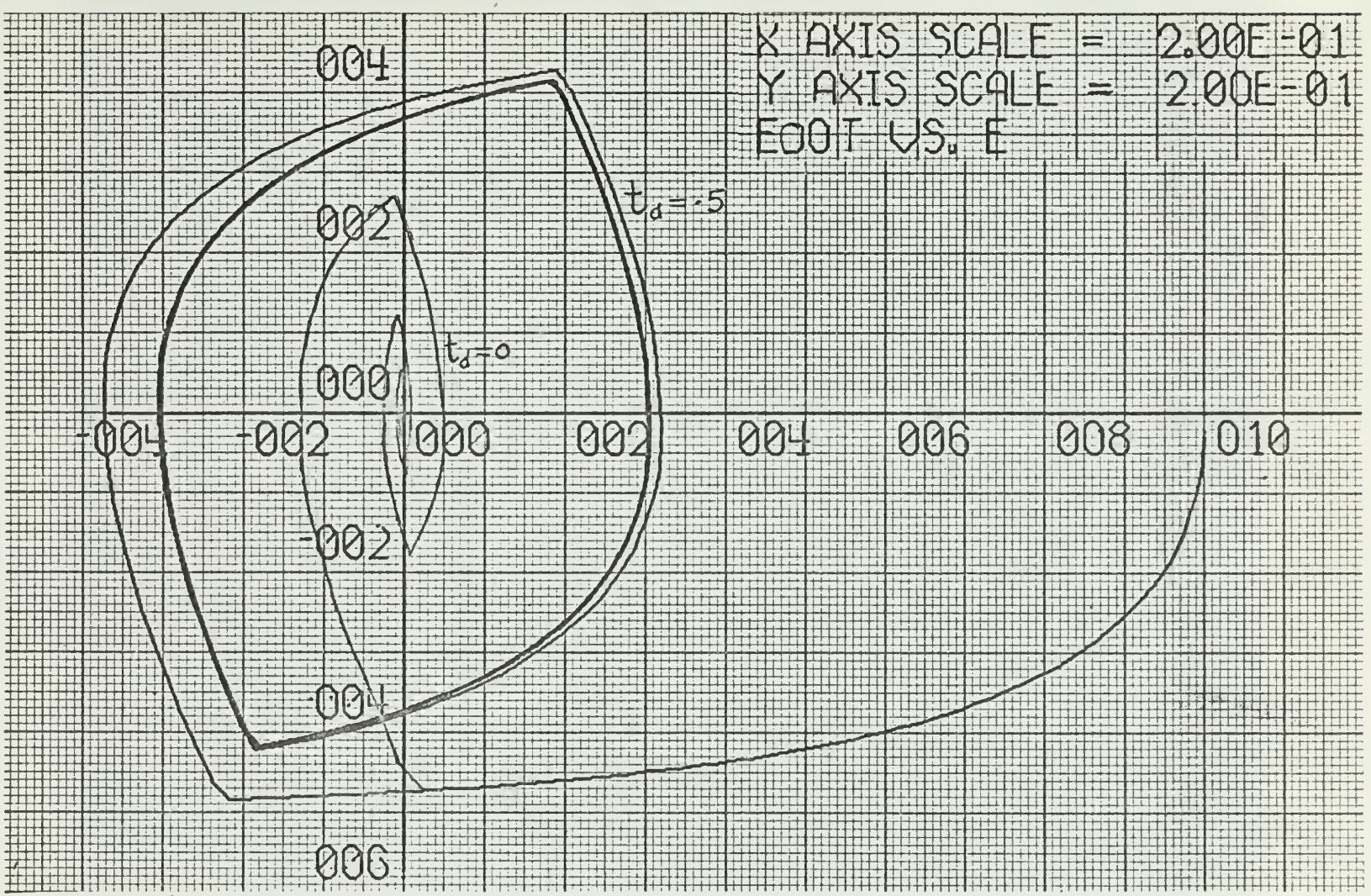


Fig. 8-11a . Phase plane diagram for a relay servo system showing the effect of time delay

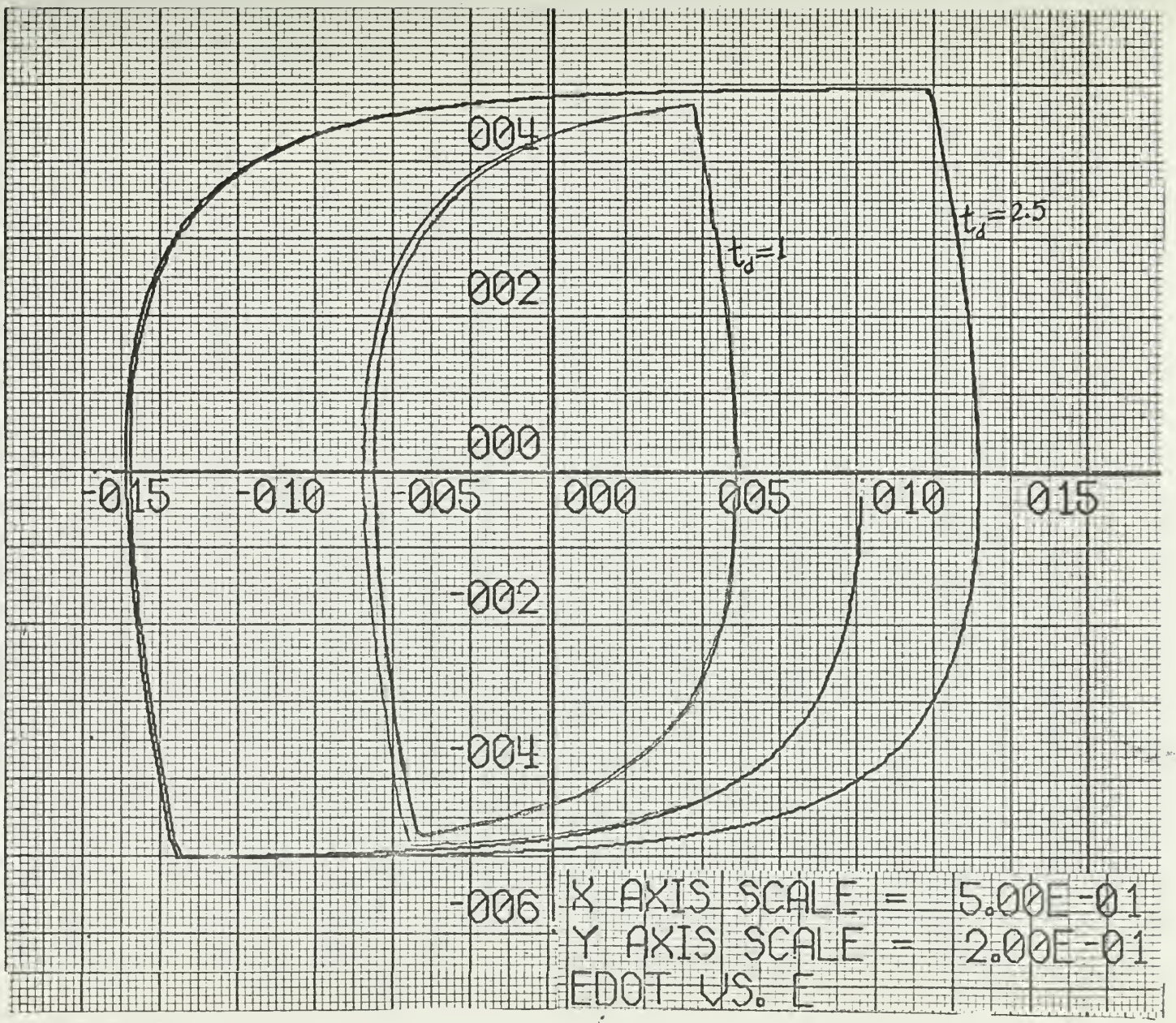


Fig. 8-11B. Phase plane diagram for a relay servo system showing the effect of time delay

is the phase plane of the relay servo system a time delay of one and 2.5 seconds plotted on the same sheet. From these phase portraits we find a most important aspect of the phase portrait is the switching line which is not vertical. The effect of the dead time is to rotate the switching line into the first and third quadrants, the result is growing oscillation.

The process of moving the switching line from vertical axis to a line in the first and third quadrants is seen to affect the nature of the response very unfavorably. On the other hand, if the switching line can be made to lie in the second and fourth quadrants by means of phase advance, or lead, on the error signal, the relay servo system is likely to be a very well behaved one, in that it has small overshoot and will reach and remain within a given value of error in a short time. This will be seen in section 9.

Dead time will result in a terminal limit cycle which may or may not be acceptably small, the amplitude of oscillation depends upon delay time as shown in Fig. 8.9 and .10.

8.6. The effect of time delay in steady state oscillation frequency.

In Fig. 8.12 the curve of system error against time is shown. Initial condition of the error is due to a suddenly applied step input signal, this gives an initial error and velocity (output rate) to the form of response waveform as shown in Fig. 8.13. After the transient has died out the output shaft follows the input by performing oscillation about a final position as shown in Fig. 8.5.

In Fig. 8.12 curves 1, 2, 3, 4, and 5 are for a delay time of .1, .25, .5, .8 and 1.0 second respectively. It indicates that the smaller time delay causes high frequency oscillation with smaller amplitude, in other words, the frequency of the oscillation is higher, and the larger value of time delay results in larger oscillation amplitude with lower frequency

as shown in Fig. 8.12.

In Fig. 8.13 curves 1, 2, 3, 4, 5, and 6 are for a delay time of 0, .1, .25, .5, .8, and 1.0 second respectively. The curves show the relation of the rate output to time delay.

From Fig. 8.5 or Fig. 8.12, the data for oscillation frequency against time delay can be obtained as shown in Table 8.2.

Table 8.2. Data for steady state oscillation frequency vs. time delay.

delay time	frequency
.1	.455
.25	.294
.5	.200
.8	.156
1.0	.135

The data plotted in log log scale sheet there appears to be a straight line as shown in Fig. 8.14; it is seen that the frequency decreases as the delay time increases. Because the curve of oscillation frequency against time delay is a straight line, therefore we can obtain an equation to express their relation as

$$f = \frac{.135}{t_d^{.57}} \qquad (8.5)$$

From Fig. 8.12 we take the steady state error for each time delay, tabulate the data as shown in Table 8.3.

Table 8.3. Steady state error vs. time delay.

Delay time	Error
0	0
.1	.09
.25	.18
.5	.31
.8	.484
1.0	.588

Fig. 8.15 is the plot of error vs. delay time. It comes out approximately a straight line, it indicates that the steady state error is proportional

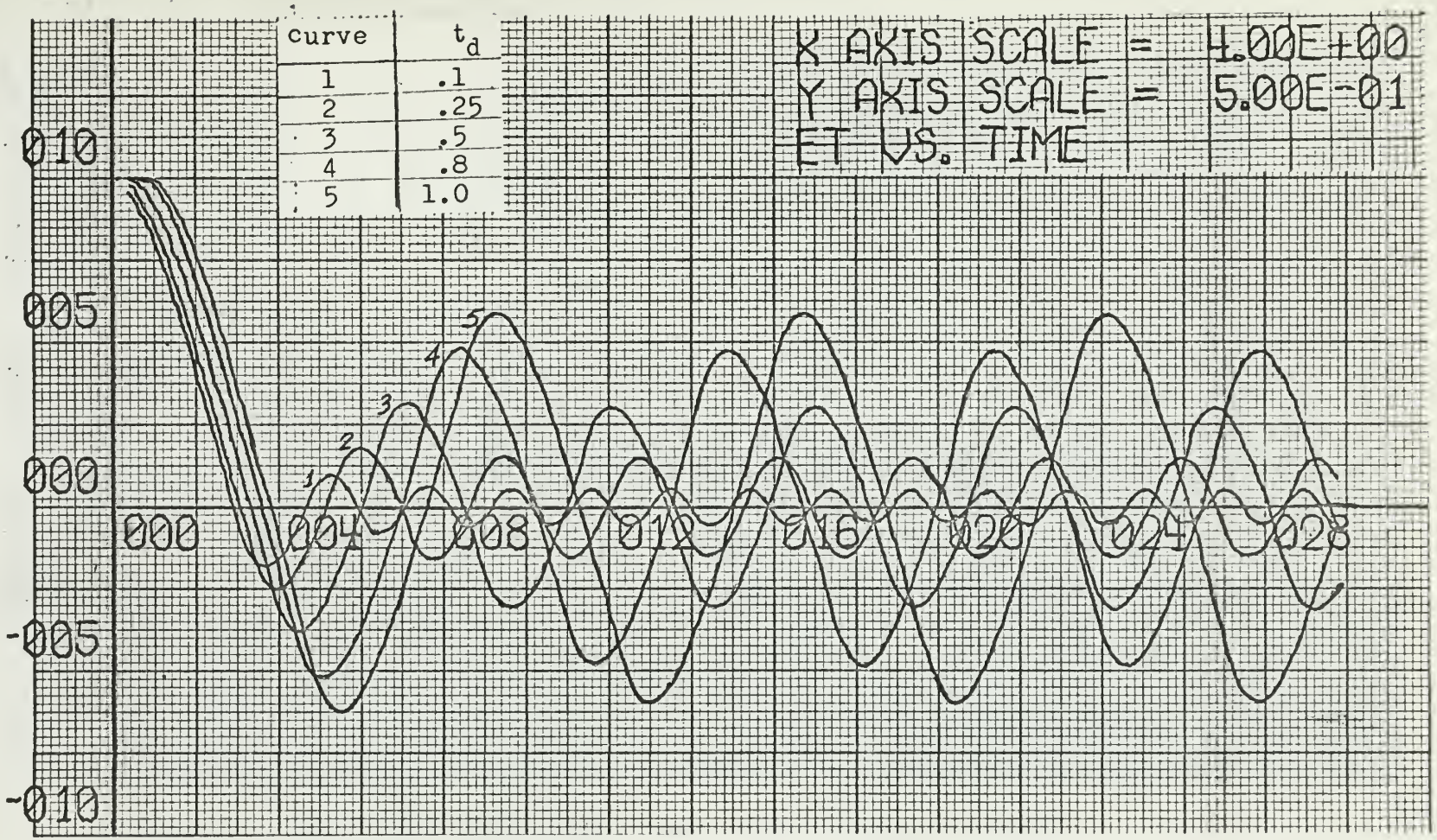


Fig. 8-12. Error vs. time curves showing the effect of time delay on the system error.

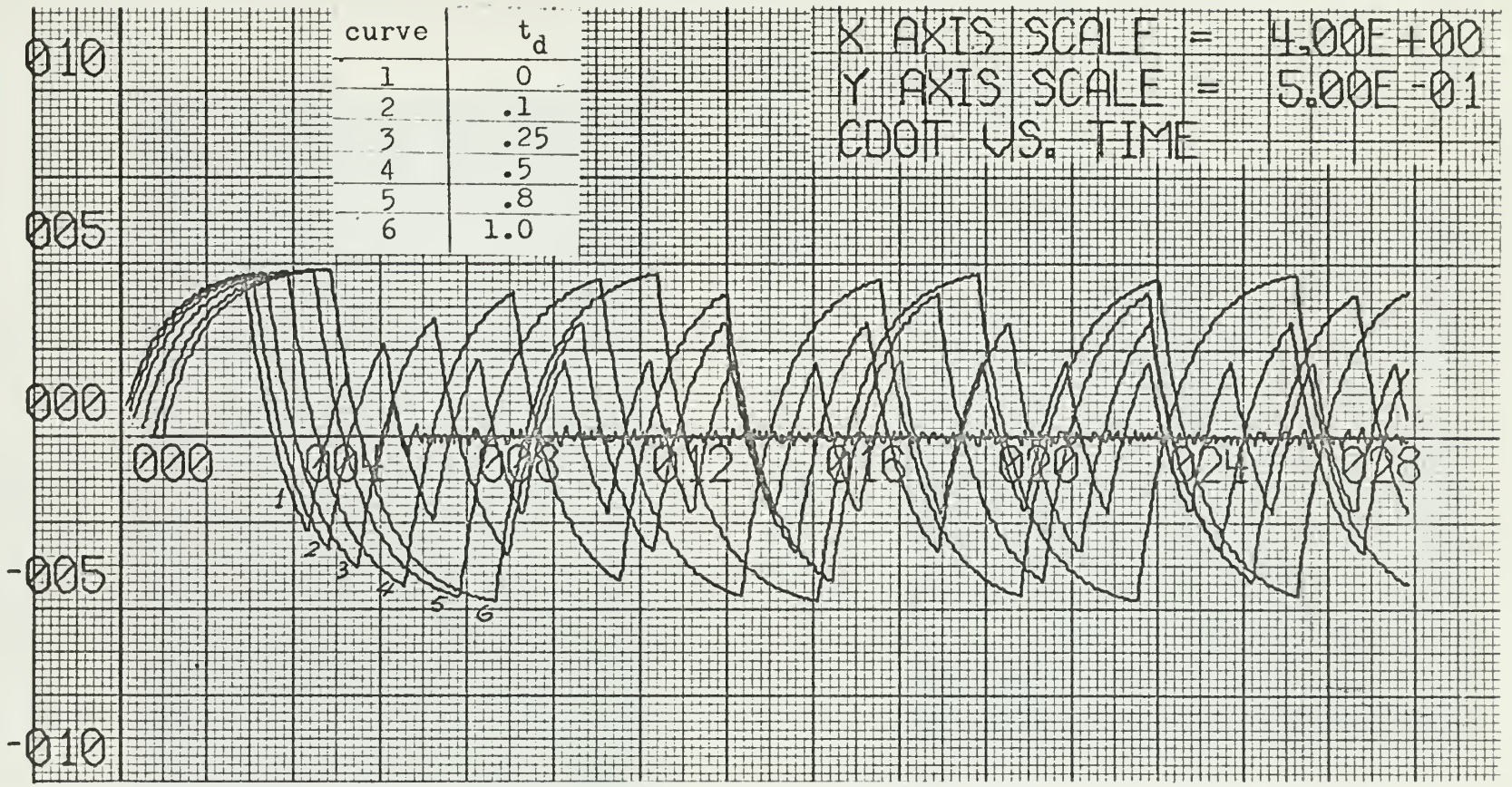


Fig. 8-13. Velocity vs.time curves showing the effect of time delay on velocity.

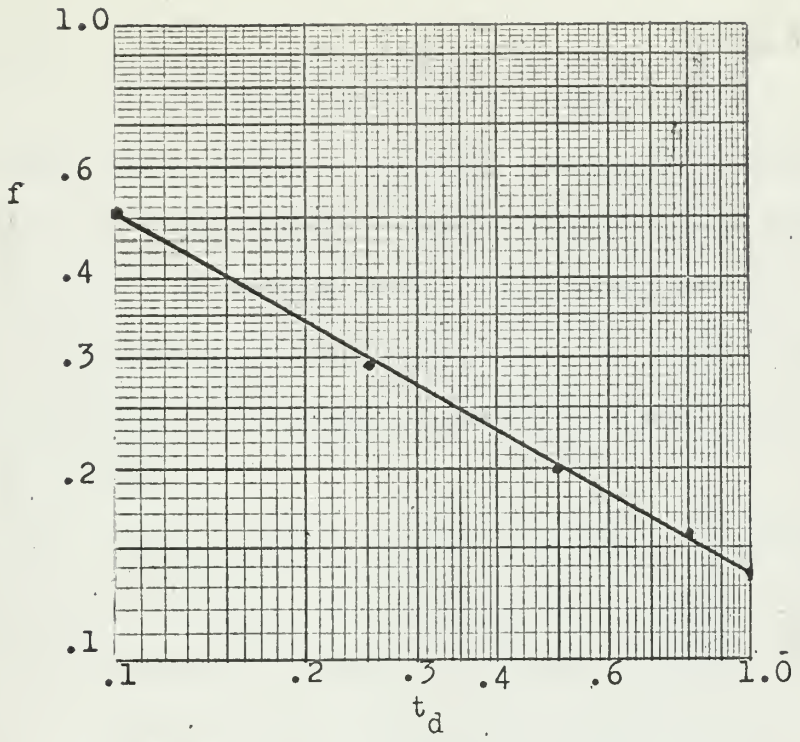


Fig 8-14. Steady state oscillation frequency vs. time delay.

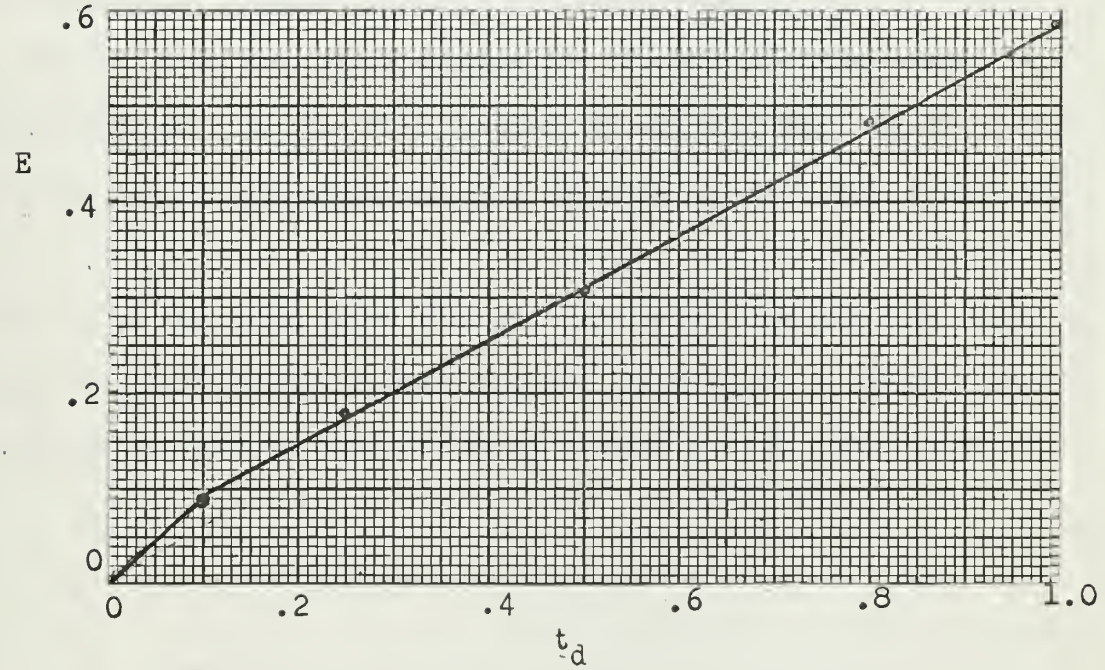


Fig. 8-15. Systemerror against time delay

to the time delay.

8.7. The effect of the magnitude of the relay voltage.

From the experimental work, we find that the magnitude of the relay voltage also has some effects on the response curve and phase trajectory of the system. When the relay voltage is smaller the response has a longer rise time and a smaller overshoot and undershoot, and there is presented some velocity saturation in the phase trajectory, the amplitude of the limit cycle is smaller.

In Fig. 8.16 shows two curves with the same delay time of .5 second and same value of k_t = .02, but with different relay voltage, curve (1) is with V =.2 volt and curve (2) is with V =.5 volt. From these curves it is indicated that the smaller relay voltage causes smaller oscillation amplitude and lower frequency, but it presents a longer time rise which is its disadvantage; so far as the time rise is concerned the larger value of relay voltage is desirable, therefore, in all the experiments we set the relay voltage equal to .5 volts. The rise time for .2 volts of relay output is 6.5 seconds, and the rise time for .5 volts is 3.35 seconds. The overshoot and undershoot for .2 volt system is 12.5 percent, and that for .5 volt is 30 percent.

In Fig. 8.17 the phase plane with the same delay time and same value of k_t is shown, but with different relay voltage, it presents a velocity saturation for the system with .2 volts of relay output, and the limit cycle is smaller.

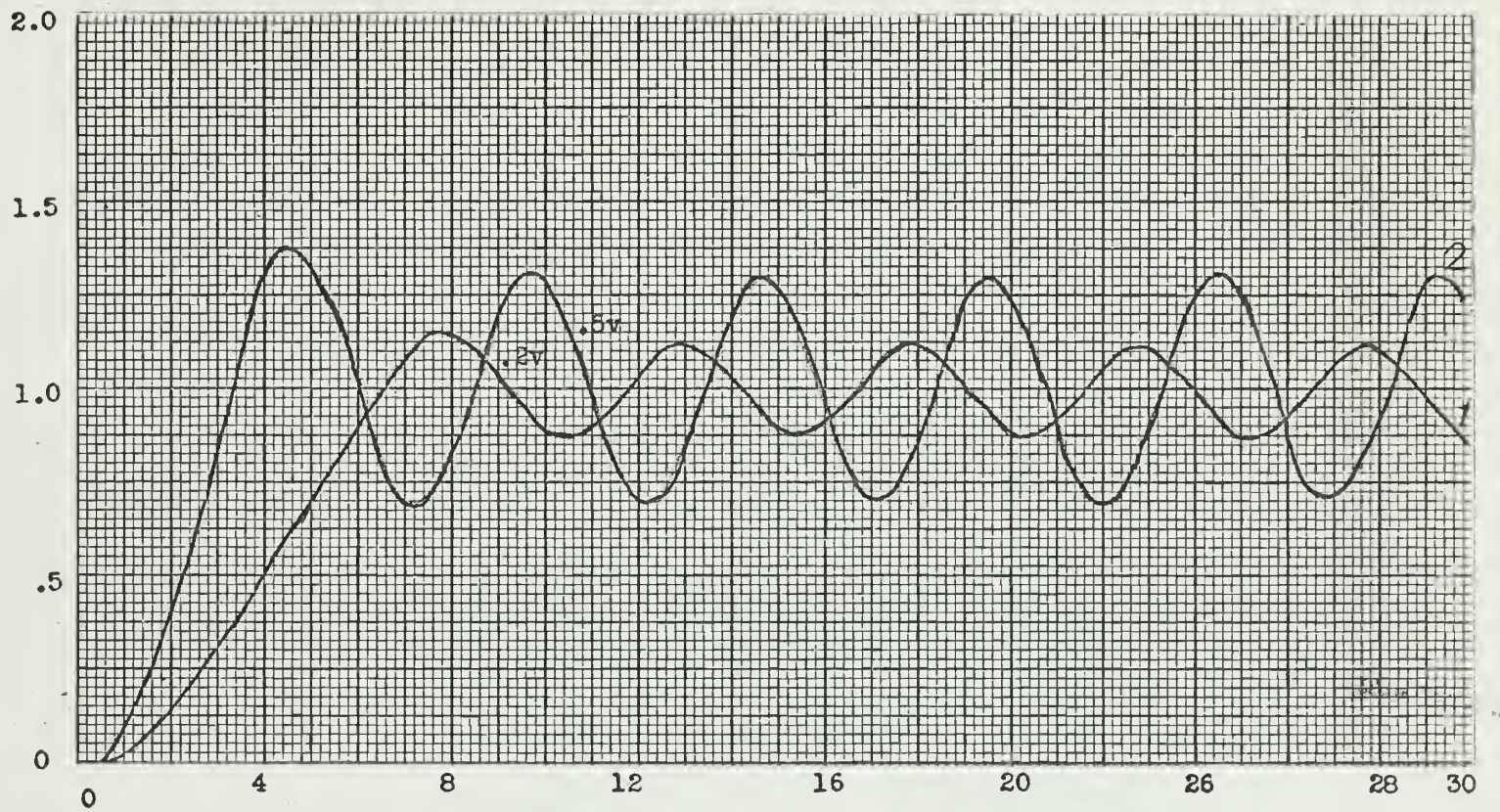


Fig. 8-16. Curves showing the effect of the relay voltage to the response. t_d=.5, k_t=.02

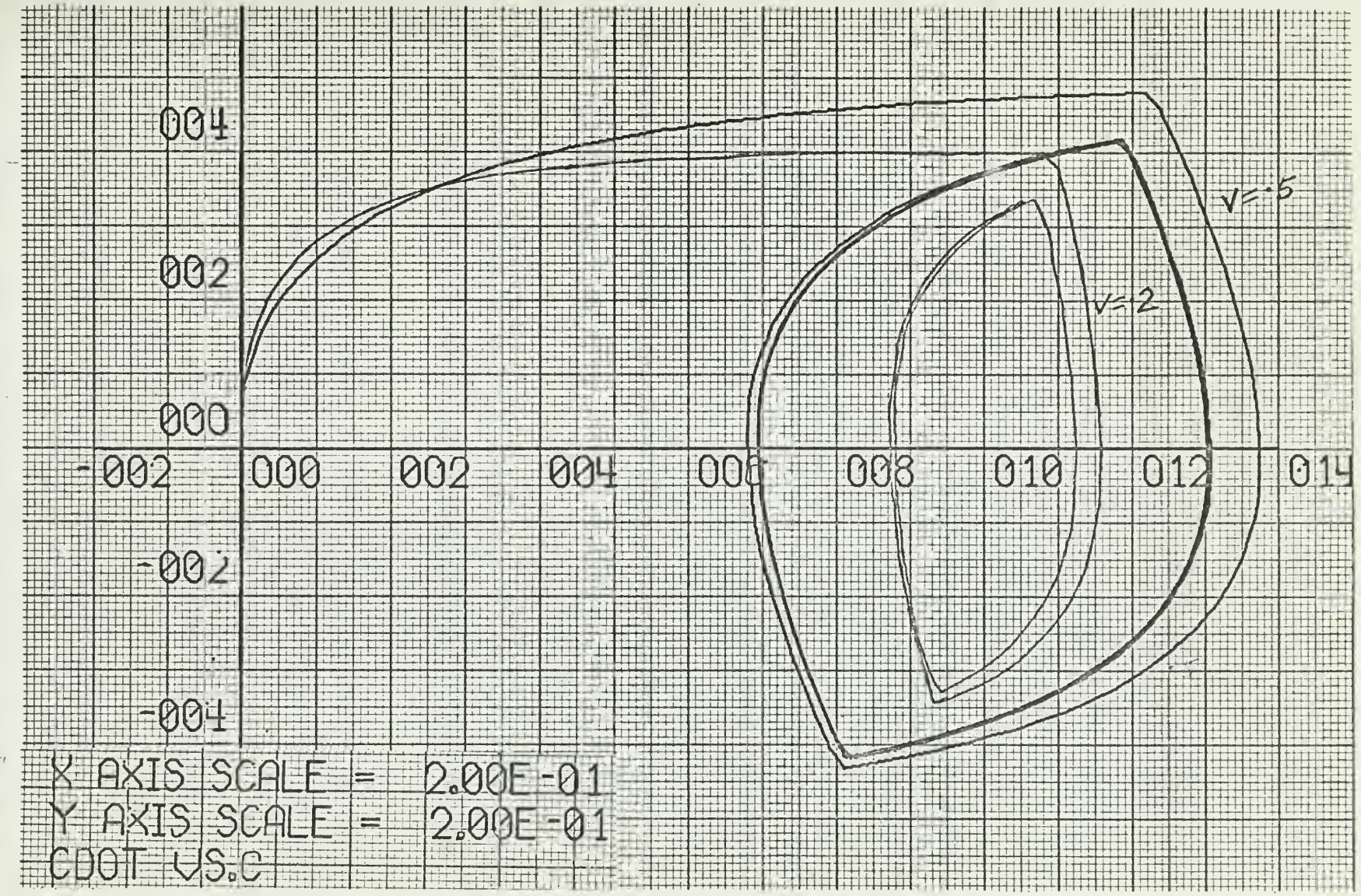


Fig. 8-17. Phase plane diagram showing the effect of relay voltege, t_d=.5 second, k_t=.02.

9. Elimination of the steady state oscillation

In many applications steady state oscillations of the output shaft are inadmissible; besides causing continual operation of relay contacts and wear of gears and bearings the power consumed in oscillations may be objectionable. The simplest method of eliminating the oscillation is to introduce a dead zone in the relays so that the error always returns to the dead zone after a disturbance. But the relay used has been thus far assumed an ideal relay with no dead zone, so we shall not discuss this possibility.

9.1 Velocity feedback stabilization

A very effective method of reducing the settling time and limit cycle oscillation of an relay-servo is to make the relay switching conditional not only on misalighment but also on its rate change. This can be achieved by feedback within the position control loop if the feedback signals can be added algebraically to the misalignment signal at the error sensing device.

The effect of negative velocity feedback is illustrated on the phase plane diagram in Fig. 9.1; and the corresponding time response is shown in Fig. 9.2.

It is obvious that the larger the value of k_t the smaller the amplitude of the limit cycle, and the more the switching line shifted to counterclockwise direction; and the amplitude of oscillation is reduced; the overshoot and undershoot for $k_t = .02$ is about 15 percent, for $k_t = .8$ is about 2 percent as shown in Fig. 9.2.

9.2 Optimising the transient response.

Because of the introduction of the time delay in the system, it is difficult to derive a method to determine the value of k_t which will give

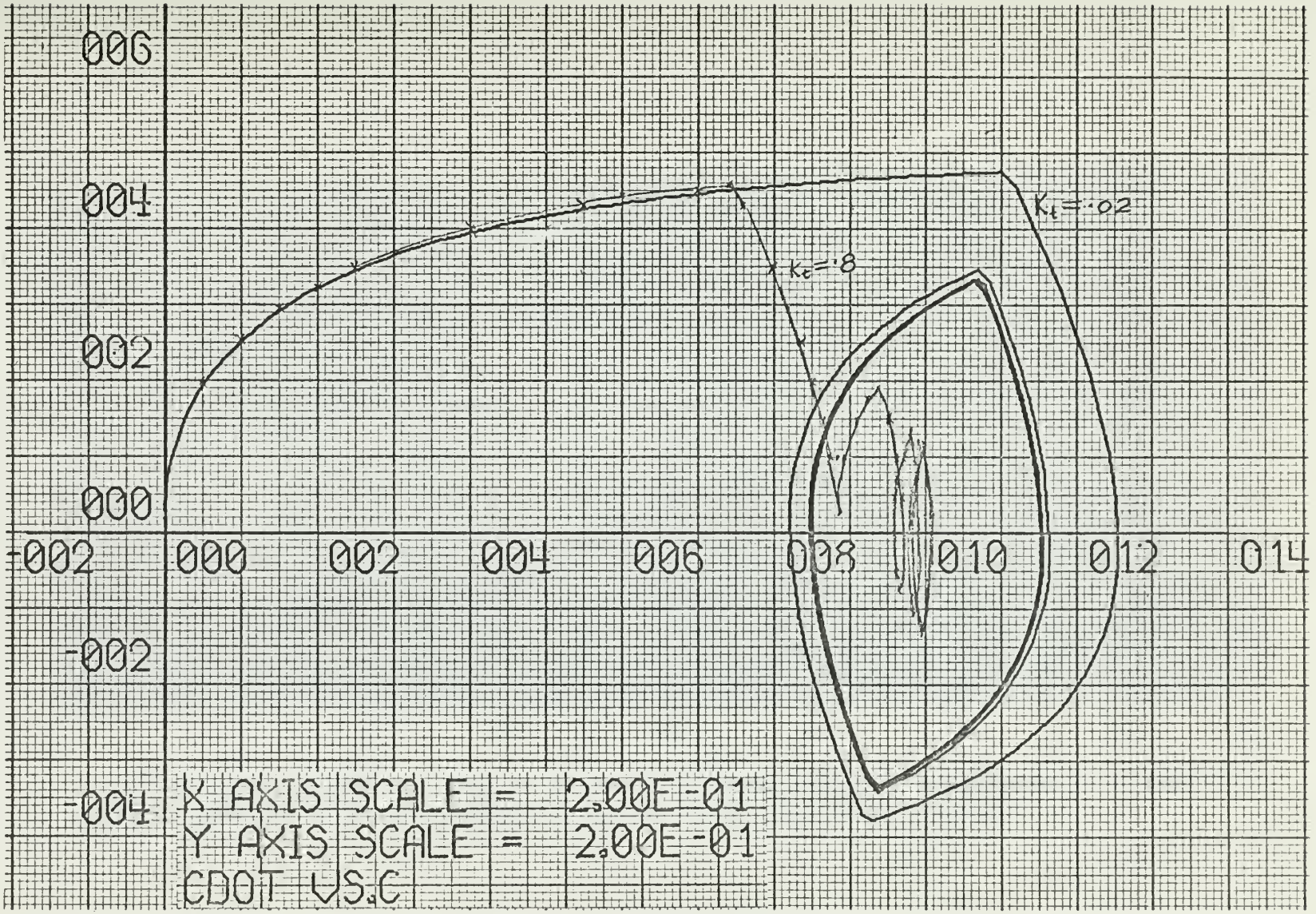


Fig. 9-1. Phase plane of relay servo system with a time delay of .25 second compensated by increasing the gain of tachometer feedback gain to .8

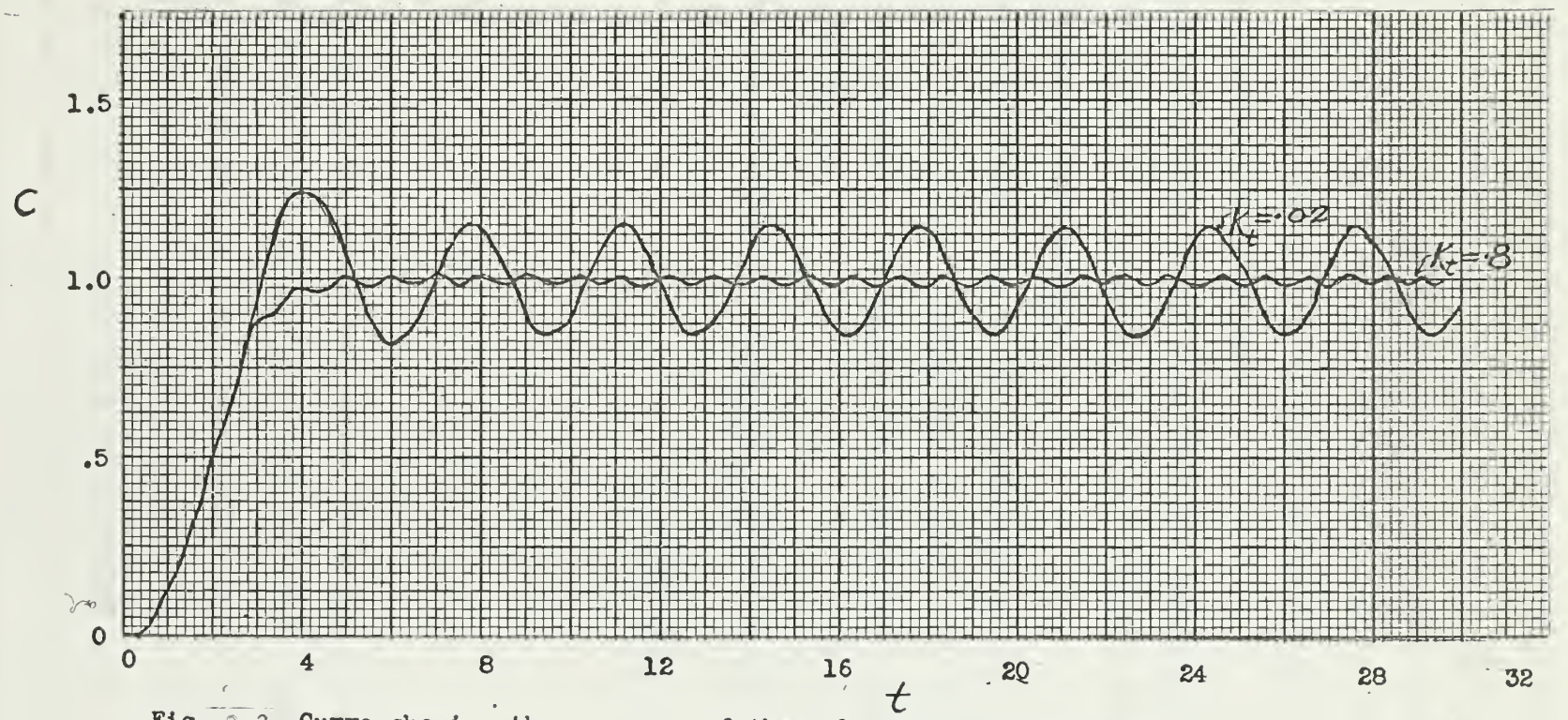


Fig. 3-2. Curve showing the response of the relay servo system be compensated to an acceptable optimum operation by in creasing k_t to a value of .8, time delay .25 second.

an optimum operation, which can only be determined through experimentation. As we know the system damping will be improved by the tachometer feedback, the net effect of the derivative signal is to cause the relay to operate early, so by using k_t we can control the switching characteristics of the relay. For many applications of relay servomechanism, the tachometer feedback compensation is adequate.

Through experimentation it was found that the system with a delay time below .5 second can be compensated for 5 percent overshoot and undershoot which is considered to be acceptable.

The response and phase plane of the compensated system for various time delays are shown in Fig. 9.3(a) through Fig. 9.3(e). They have been compensated by velocity feedback. The values of k_t (which give acceptable optimum operation) are as follows:

Table 9.1 Optimum value of k_t for eliminating the effect of time delay.

t_d	k_t	o/o overshoot and undershoot
0	.3	/
.1	.4	.1
.25	.8	1.6
.5	1.5	5.2

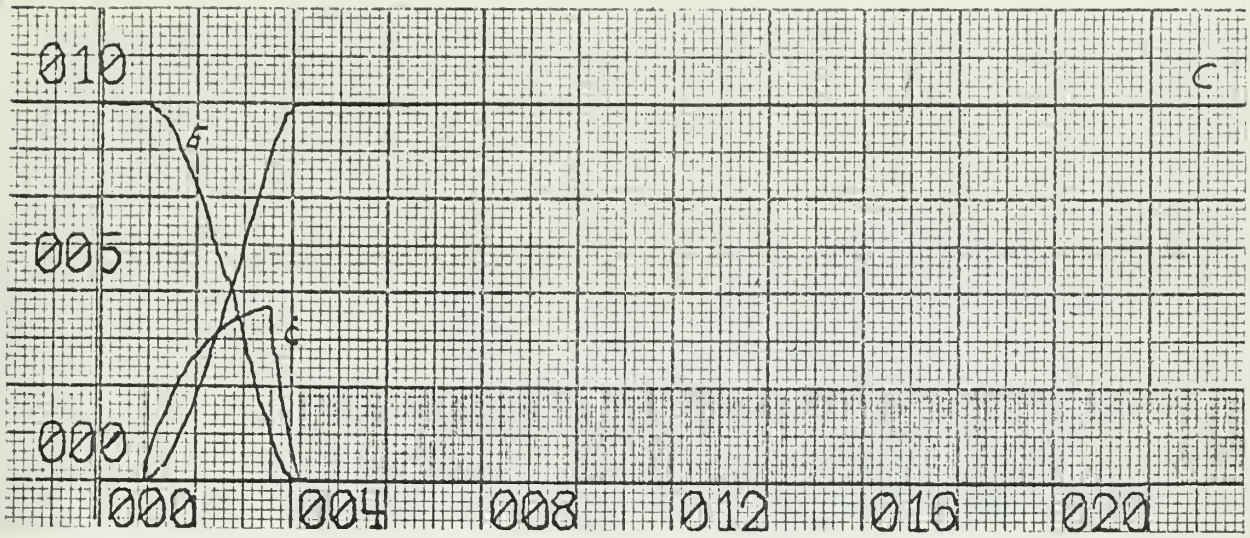
For systems with .8 and 1.0 second time delay, it is found that they can not be compensated for 5 percent overshoot and undershoot, even when we increase the tachometer gain, k_t, to a considerable value, the amplitude of steady state oscillation is still rather large as shown in Fig. 9.4 and Fig. 9.5; the only effect of increasing k_t at this time is to shift the whole response curve downward, as shown in Fig. 9.5 the response of the system with one second time delay at $k_t = 4.5$ is rather low but still with

considerable oscillation amplitude which is almost the same as that when k_t = 1.5, which is beyond the acceptable value.

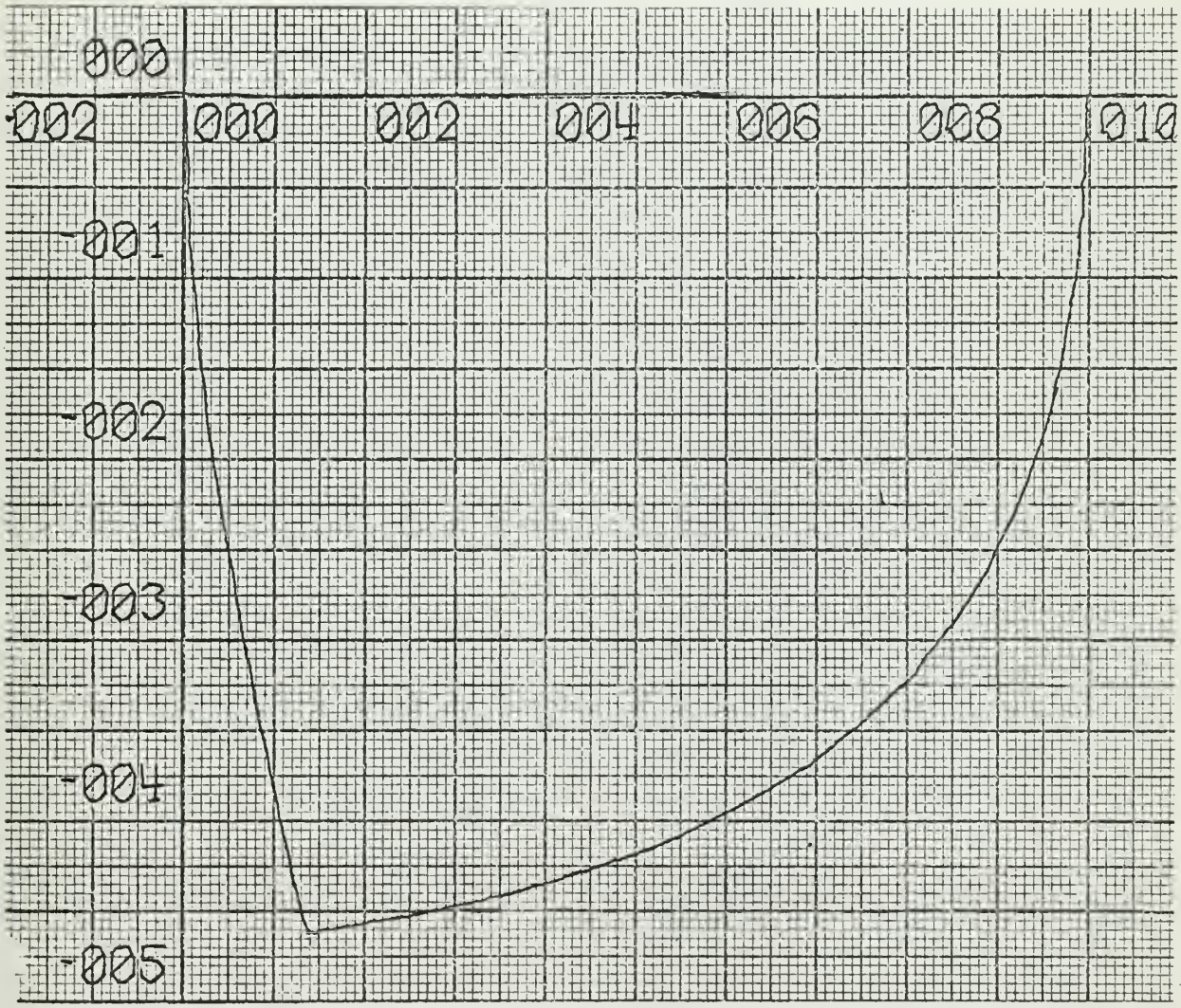
Fig. 9.6 illustrates the time delay vs. k_t curve at the range from 0 to .5 second time delay. The curve appears to be a straight line, but when the time delay above .5 second occurs this situation does not hold, since the system with a delay time above .5 second cannot be compensated. When the time delay is less than .5 seconds, the system can be compensated; and because t_d vs. k_t curve is a straight line, therefore, it can be expressed by an equation as given in equation 9.1.

$$K_t = .12 + 2.8\, t_d \qquad (9.1)$$

Table 9.2 and 9.3 are the results for comparing the uncompensated and compensated system. The data in Table 9.2 is for the undelayed system for which k_t = .3 will compensate the system to give optimum operation. The data in Table 9.3 is for a .5 second time delayed system for which k_t = 1.5 will compensate the system to give acceptable optimum operation.

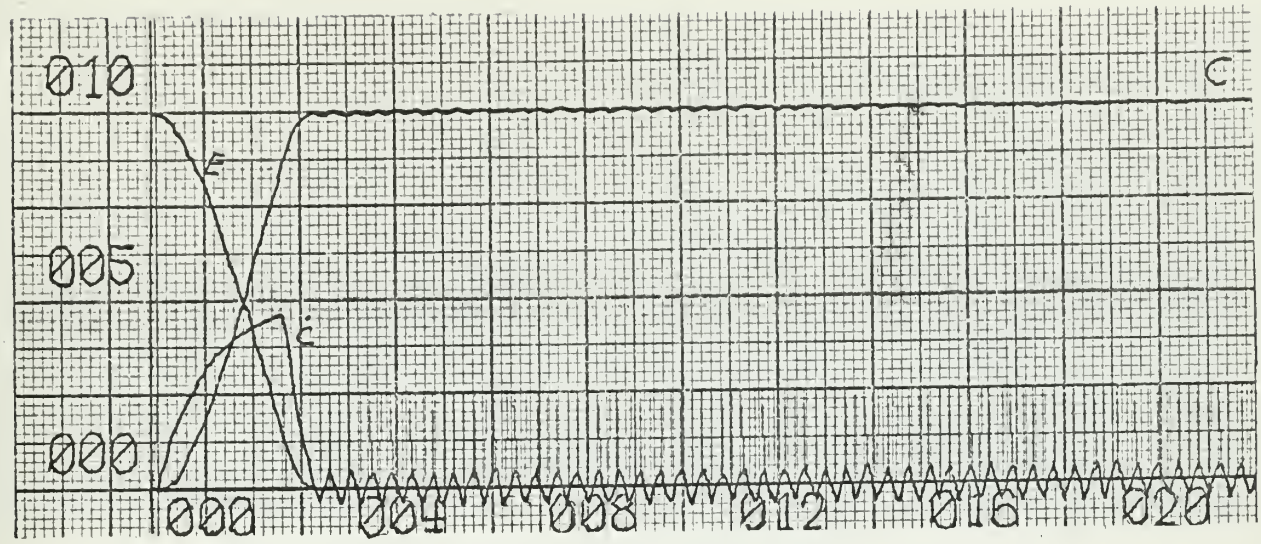


plant output, velocity and error vs. time

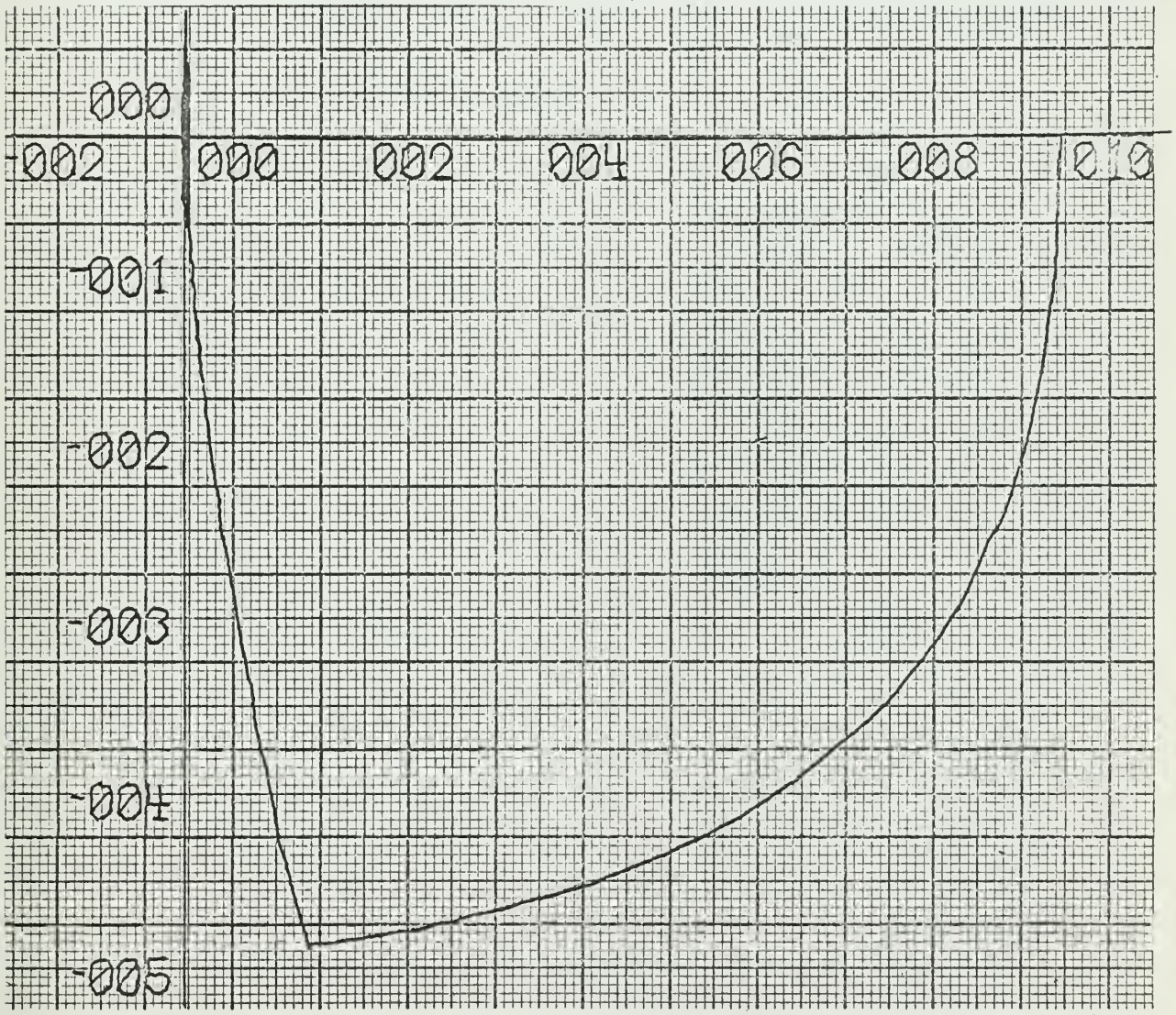


phase plane (E vs E)

Fig. 9-3a. Response and phase plane, $t_d = 0$, $k_t = .3$

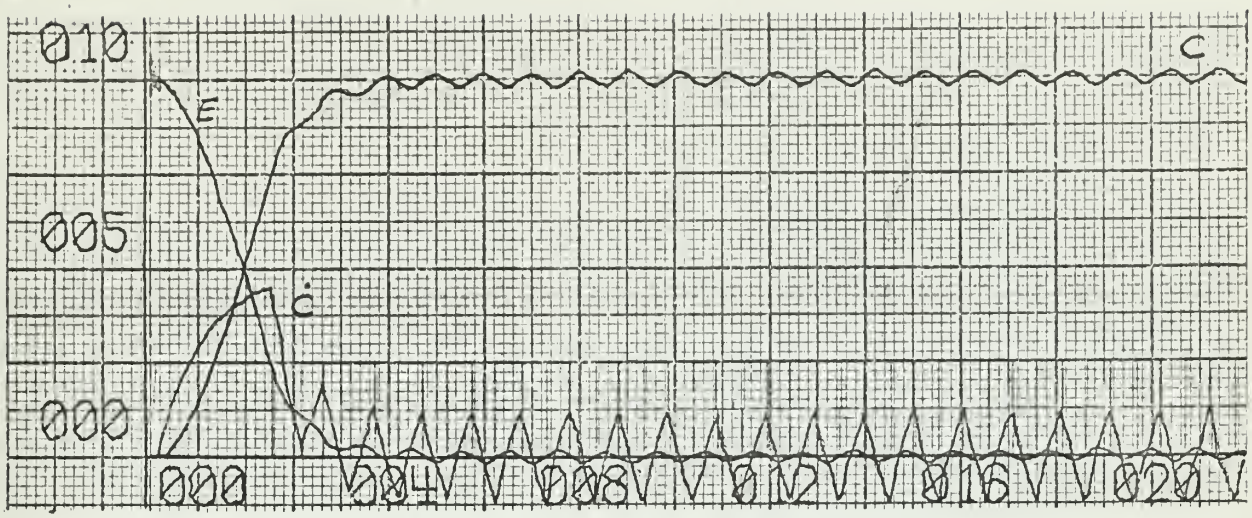


plant output, velocity and error vs. time



Phase plane

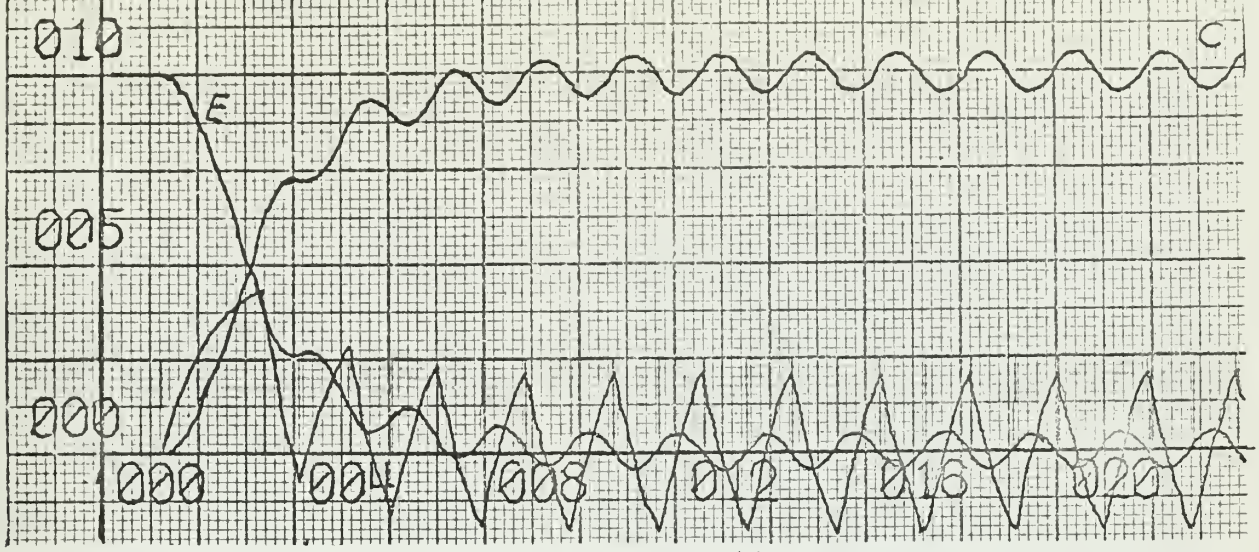
Fig. 9-3b. Response and phase plane, $t_d = .1$ sec. $k_t = .4$



plant output, velocity and error vs. time

phase plane

Fig. 9-3c. Response and phase plane, t_d = .25 sec. k_t = .8



plant output, velocity and error vs. time

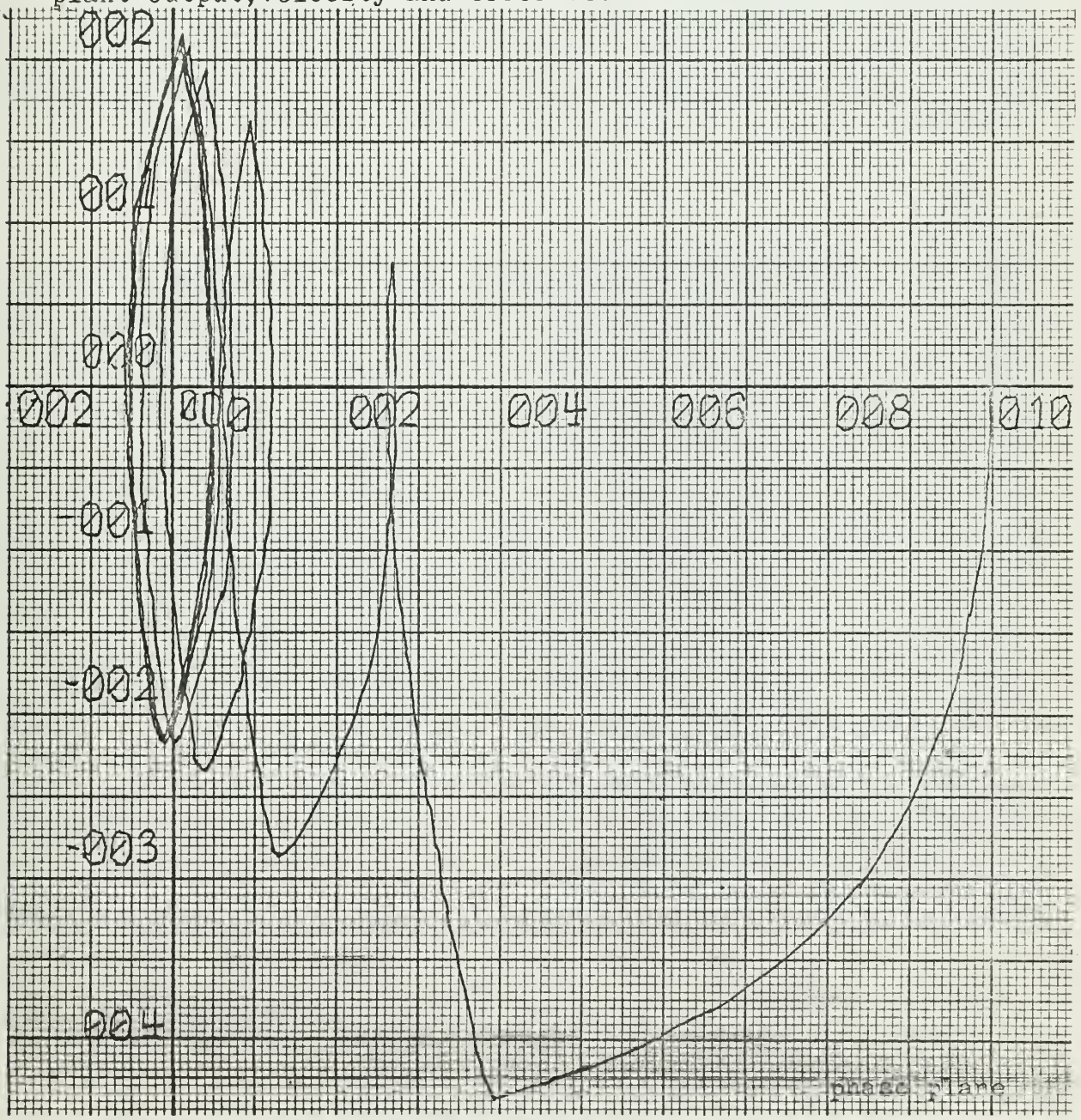
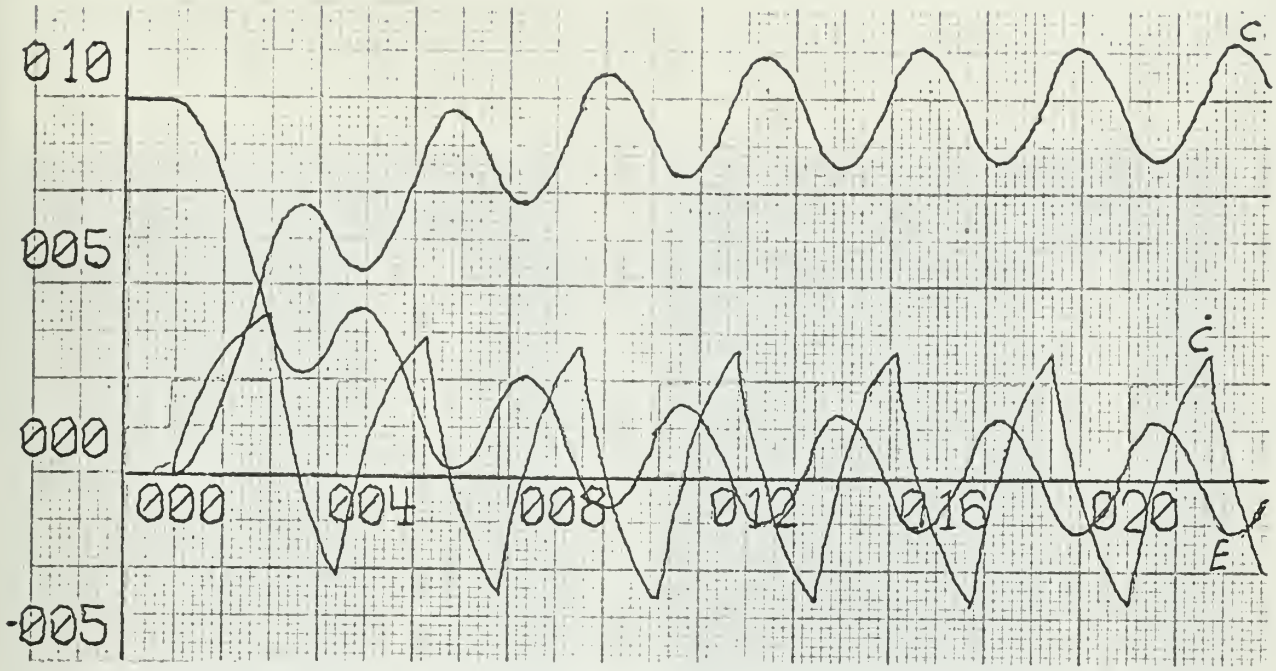
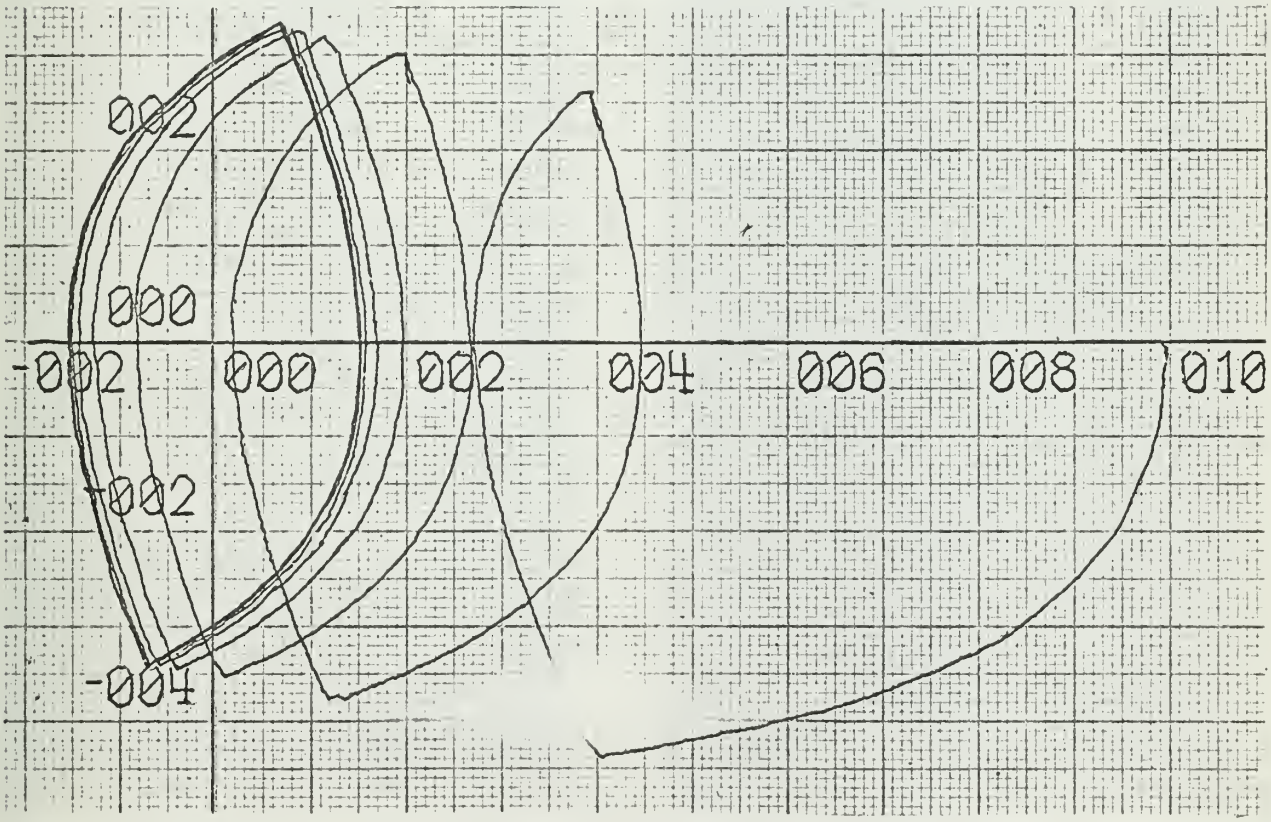


Fig. 9-3d. REsponse and phase plane, t_d = .5 sec. k_t = 1.5



plant output, velocity and error vs. time



phase plane

Fig. 9-3e. Response and phase plane, $t_d = 1.0$ sec. $k_t = 2.4$

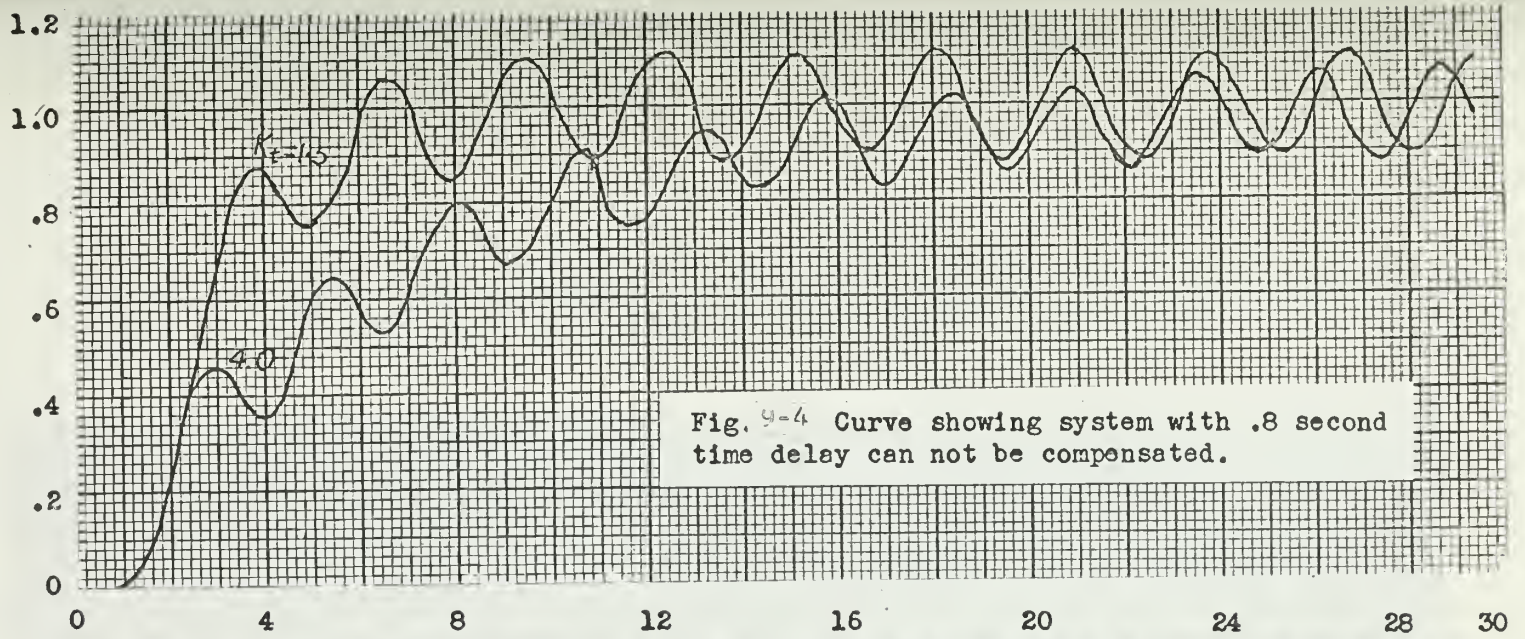


Fig. 9-4 Curve showing system with .8 second time delay can not be compensated.

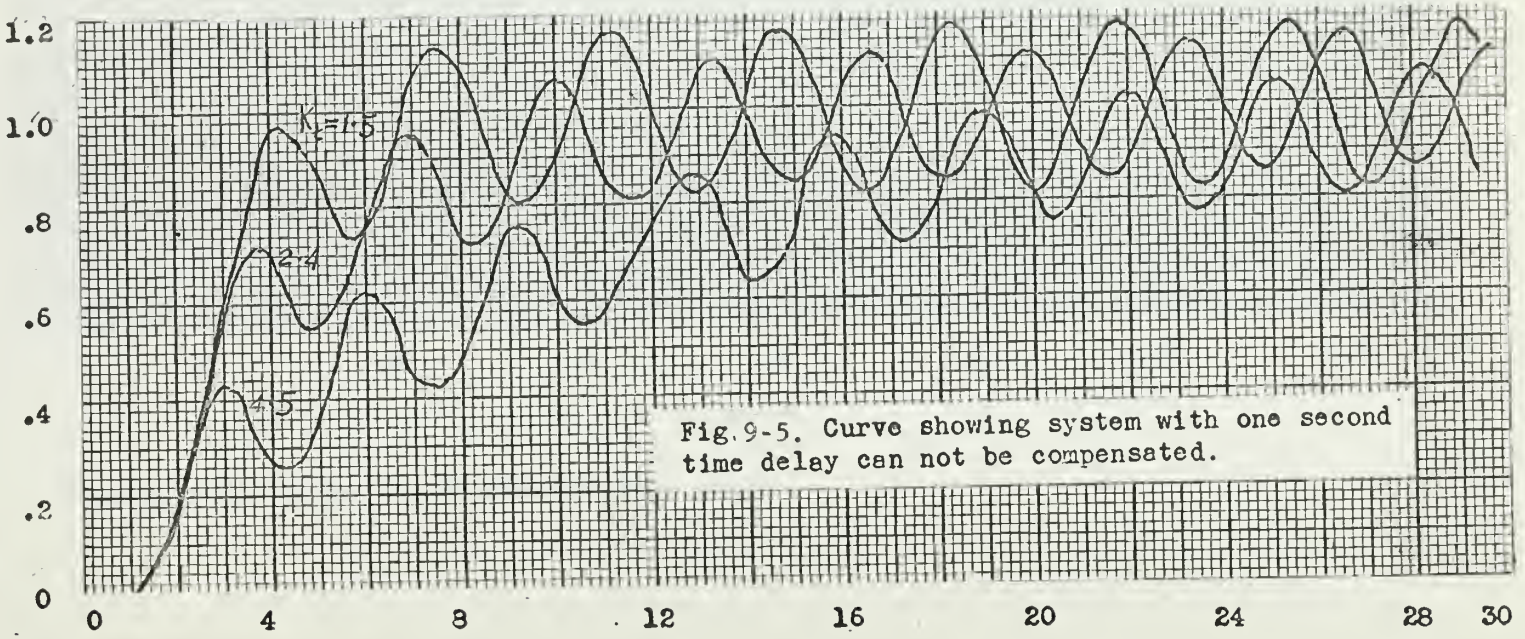


Fig. 9-5. Curve showing system with one second time delay can not be compensated.

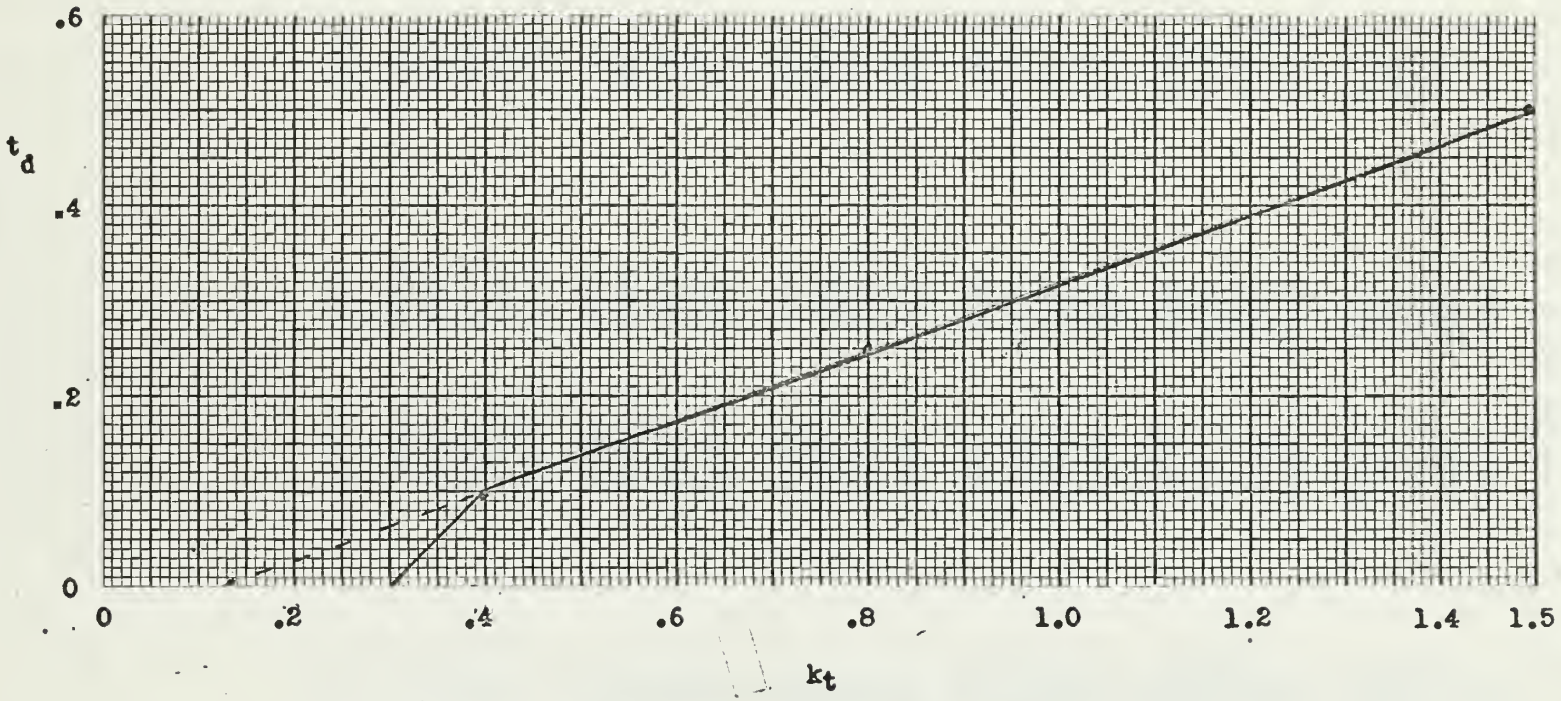


Fig. 9-6. Time delay vs. tachometer gain curve

Table 9.2 Calculated response of relay servo system Forcing function = unit step t_d = 0 second

Time	System error		Rate error		Plant output	
	k_t = .02	k_t = .3	k_t = .02	k_t = .3	k_t = .02	k_t = .3
0	1.0000	1.0000	.0000	.0000	.0000	.0000
.5	.9467	.9467	-.1967	-.1967	.0531	.0531
1.0	.8161	.8161	-.3161	-.3161	.1839	.1839
1.5	.6384	.6384	-.3884	-.3884	.3616	.3616
2.0	.4323	.4323	-.4323	-.4323	.5677	.5677
2.5	.2090	.2090	-.4590	-.4590	.7910	.7910
3.0	-.0223	.0296	-.4052	-.1802	1.0223	.9704
3.5	-.1285	.0023	-.0490	-.0124	1.1285	.9977
4.0	-.0945	.0006	.1670	-.0016	1.0945	.9994
4.5	.0192	.0003	.1988	-.0005	.9808	.9997
5.0	.0442	.0002	-.0761	-.0003	.9558	.9998
5.5	-.0206	.0002	-.0629	-.0002	1.0206	.9998
6.0	-.0054	.0002	.0879	-.0002	.9946	.9998
6.5	-.0056	.0002	-.0232	-.0002	1.0056	.9998
7.0	-.0005	.0002	.0156	-.0002	1.0005	.9998
7.5	-.0006	.0002	-.0046	-.0002	1.0000	.9998
8.0	-.0006	.0002	-.0045	-.0002	1.0006	.9998
8.5	.0002	.0002	-.0056	-.0002	.9998	.9998
9.0	.0003	.0002	.0217	-.0002	.9997	.9998
9.5	-.0002	.0002	.0042	-.0002	1.0002	.9998
10.0	.0002	.0002	.0106	-.0002	.9998	.9998
10.5	-.0002	.0002	.0115	-.0002	1.0002	.9998
11.0	.0007	.0002	.0054	-.0002	.9993	.9998
11.5	-.0001	.0002	-.0066	-.0002	1.0001	.9998
12.0	.0004	.0002	.0058	-.0002	.9996	.9998
12.5	-.0003	.0002	.0179	-.0002	1.0003	.9998
13.0	-.0009	.0002	-.0342	-.0002	1.0009	.9998
13.5	-.0000	.0002	.0128	-.0002	1.0000	.9998
14.0	-.0004	.0002	-.0067	-.0002	1.0004	.9998
14.5	-.0004	.0002	.0206	-.0002	1.0004	.9998
15.0	-.0000	.0002	-.0006	-.0002	1.0000	.9998

response of relay servo system Forcing function = unit step $t_1 = .5$ second

Time	System error		Rate error		Plant output	
	$k_t = .02$	$k_t = 1.5$	$k_t = .02$	$k_t = 1.5$	$k_t = .02$	$k_t = 1.5$
0	1.0000	1.0000	.0000	.0000	.0000	.0000
.5	1.0000	1.0000	.0000	.0000	.0000	.0000
1.0	.8467	.8467	-.1967	-.1967	.0533	.0533
1.5	.6161	.6161	-.3161	-.3161	.1839	.1839
2.0	.4384	.4384	-.3884	-.3884	.3616	.3616
2.5	.3323	.3323	-.4323	-.4323	.5677	.5677
3.0	.2790	.2790	-.4590	-.4299	.7210	.7210
3.5	-.0249	.2762	-.4751	.0217	1.0249	.7238
4.0	-.2625	.2315	-.4150	-.1836	1.2625	.7685
4.5	-.3726	.1080	-.0550	.2440	1.3726	.8920
5.0	-.3409	.0653	.1634	-.0487	1.3409	.9347
5.5	-.2234	.1217	.2958	.0574	1.2234	.8783
6.0	-.0537	.0910	.3762	-.1619	1.0537	.9090
6.5	.1476	-.0034	.4249	-.0972	.8524	1.0034
7.0	.3401	.0116	.1381	.1376	.6959	.9884
7.5	.3051	.0766	-.1130	.0162	.6949	.9234
8.0	.2074	.0297	-.2653	-.1869	.7926	.9703
8.5	.0498	-.0383	-.3576	-.0051	.9502	1.0383
9.0	-.1442	.0129	-.4136	.1936	1.1442	.9871
9.5	-.2950	.0590	-.1208	-.0407	1.2950	.9410
10.0	-.2923	-.0103	.1188	-.2187	1.2923	1.0103
10.5	-.1923	-.0430	.2688	.0641	1.1923	1.0430
11.0	-.0332	.0329	.3598	.1651	1.0332	.9671
11.5	.1616	.0446	.4149	-.0906	.8384	.9554
12.0	.2979	-.0374	.0994	-.1247	.7021	1.0374
12.5	.2838	-.0332	-.1364	.1211	.7162	1.0332
13.0	.1768	.0474	-.2795	.0817	.8232	.9526
13.5	.0136	.0263	-.3662	-.1472	.9864	.9737
14.0	-.1838	-.0500	-.4189	.0444	1.1838	1.0500
14.5	-.3019	-.0142	-.0676	.1698	1.3019	1.0142
15.0	-.2752	.0527	.1557	-.0080	1.2752	.9473

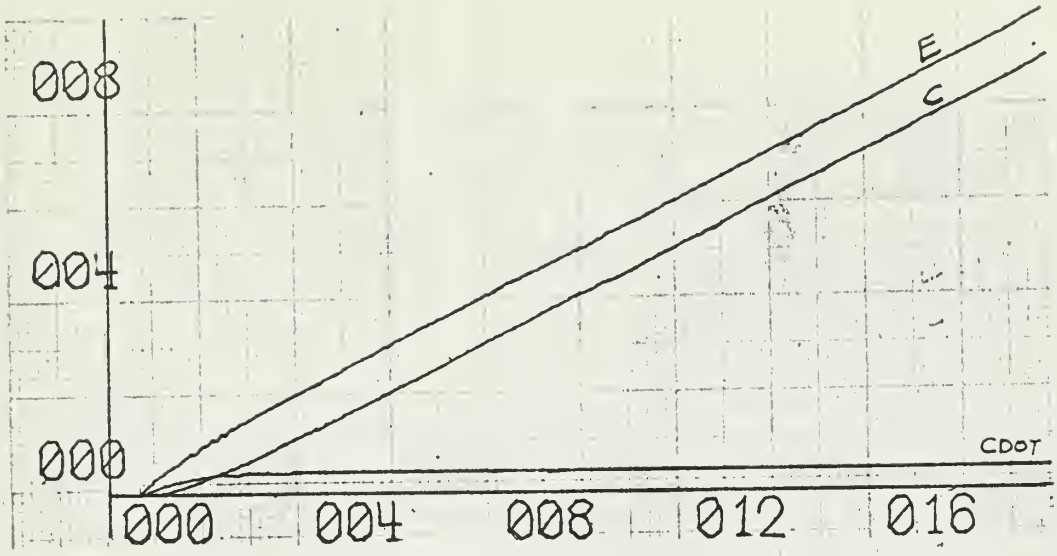
10. The response of relay servo system to ramp function input.

In the previous sections we have been concerned primarily with the response to step function input. Experimental results obtained from the relay servo system with a ramp function input are shown in several graphs.

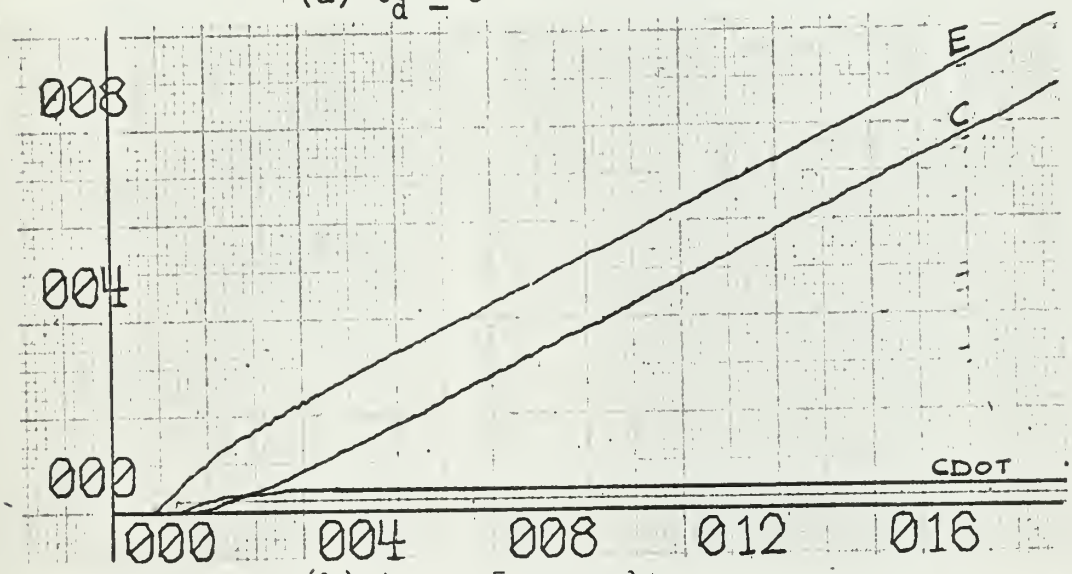
When a unit ramp input is applied to the relay servo system, the output and the error of the system are shown as a straight line, and the velocity for most part is a horizontal line, as shown in Fig. 10.1. The Cdot vs. C and Edot vs. E phase planes have the same shape, merely the Cdot vs. C plane appears above X- axis. Fig. 10.2 shows the Edot vs. E plane. This is with a relay voltage equal to .5 volt. When the relay voltage is increased to 1.0 volt the same situation remains. For relay voltage at .5 and 1.0 volt the only effect of the time delay on the relay servo system is to cause some delay in the response, no oscillation has been observed.

But when the relay voltage is increased to 1.5 volts, this situation is changed. For an undelayed system the output still appears as a straight line but the system error is no more a straight line, and the velocity curve is with some oscillations as shown in Fig. 10.3. Cdot vs. C and Edot vs. E phase planes have a different shape as shown in Fig. 10.4 and Fig. 10.5. For delayed system the output no more remains a straight line but with some oscillations, and the system error and velocity are much different than that of the system with relay voltage of .5 or 1.0 volt as shown in Fig. 10.6. The Cdot vs. C and Edot vs. E phase planes are quite different as shown in Fig. 10.7 and Fig. 10.8 respectively. Oscillations and limit cycle has been observed when the relay voltage is 1.5 volts.

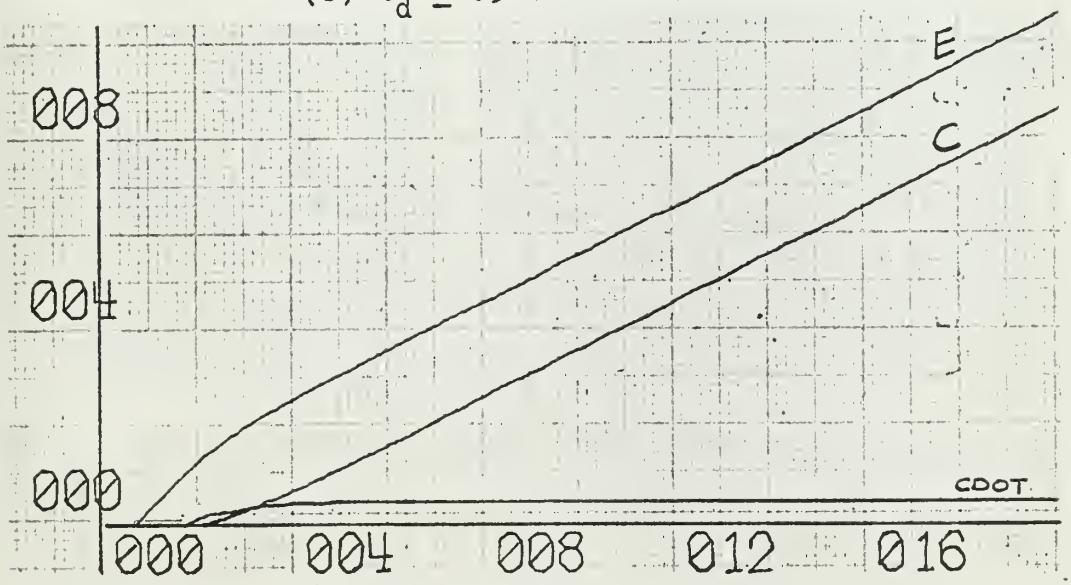
The effect of time delay on the response of the system can also be compensated by velocity feedback. For the .5 second time delayed system, when $k_t = 2.0$, the response can be compensated to an acceptable optimum operation as shown in Figures 10.9,10, 11.



(a) $t_d = 0$



(b) $t_d = .5$ second



(c) $t_d = 1.0$ second

Fig. 10-1. C, CDOT, E vs. time

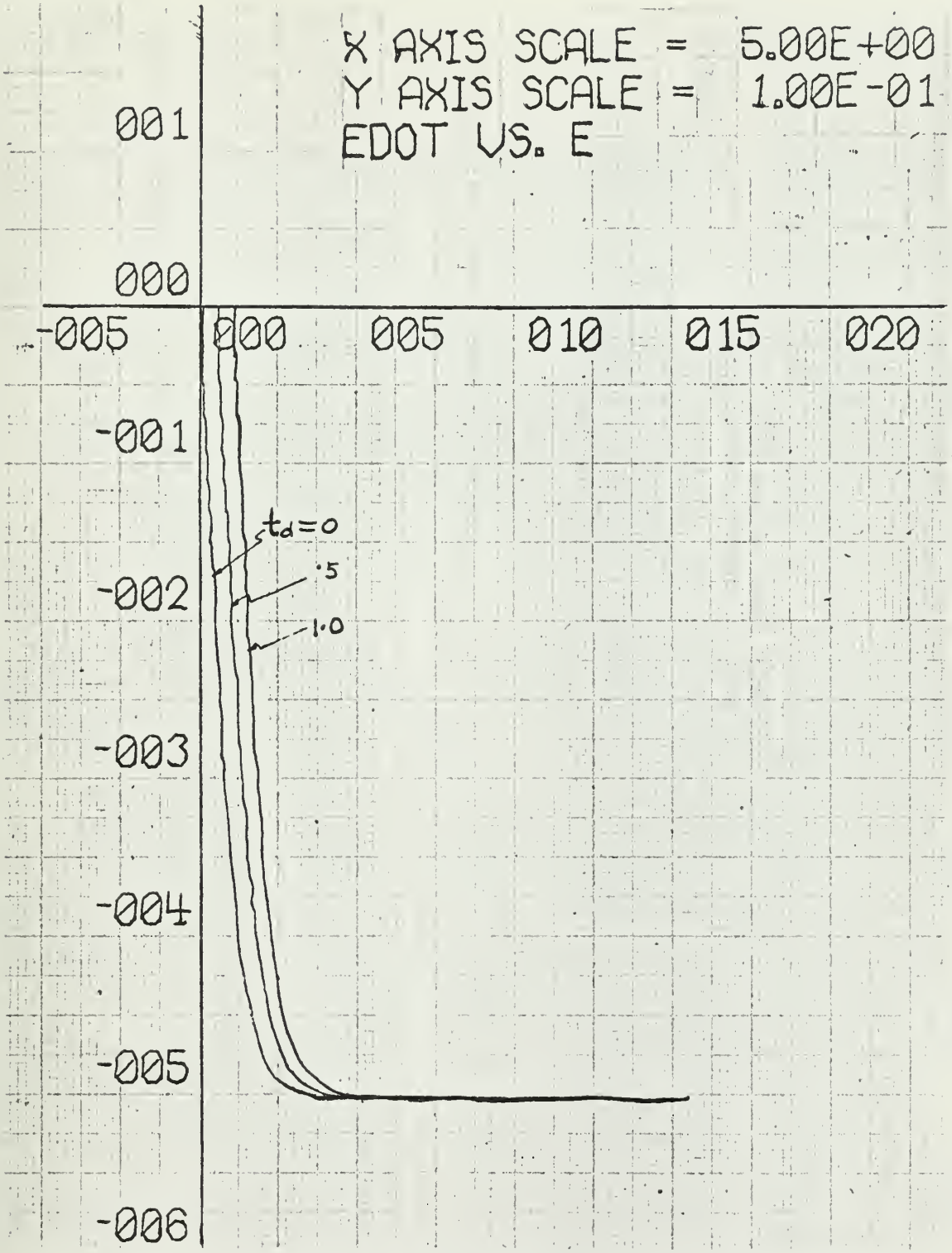


Fig. 10-2. Phase trajectory of relay servo system, relay voltage .5 volt

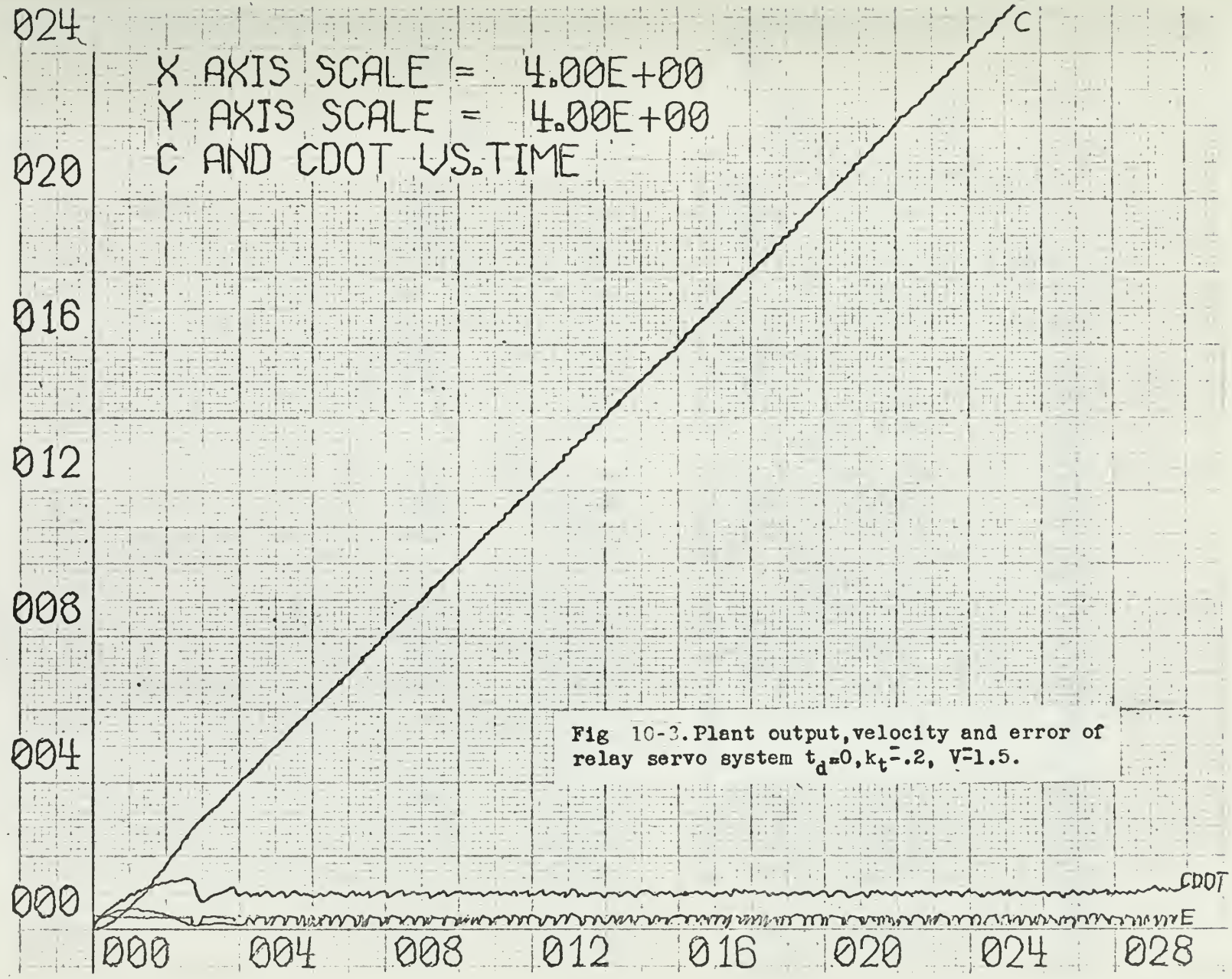


Fig 10-3. Plant output, velocity and error of relay servo system $t_d=0, k_t=.2, V=1.5$.

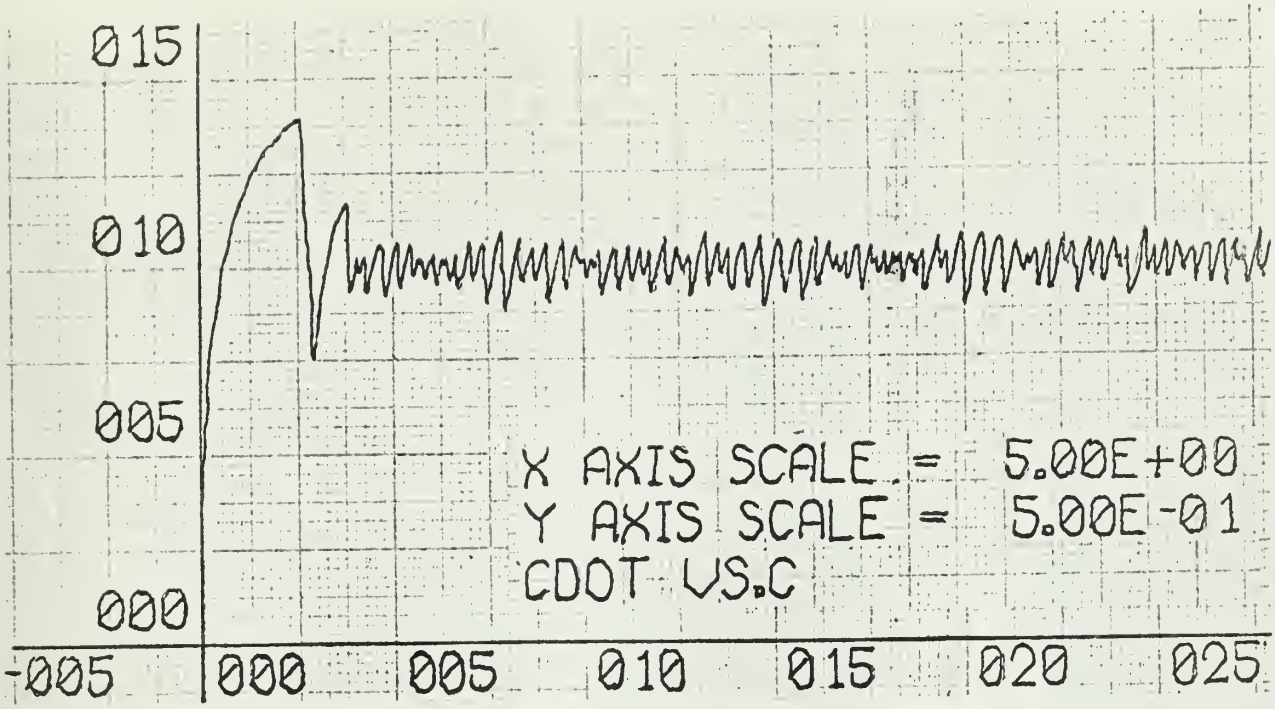


Fig. 10-4. CDOT vs C phase plane of a system with 0 time delay, k_t=.02

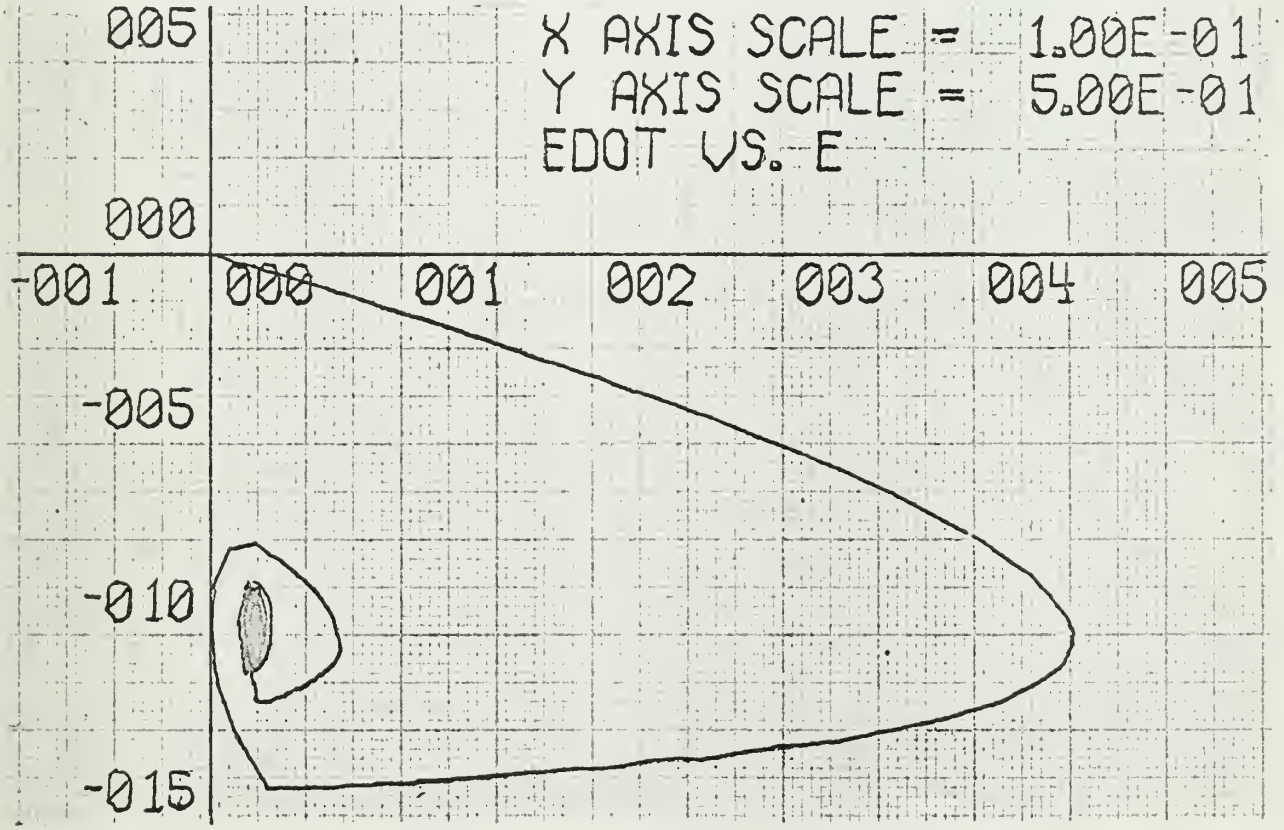


Fig. 10-5. EDOT vs. E phase plane , t_d=0, k_t=.02

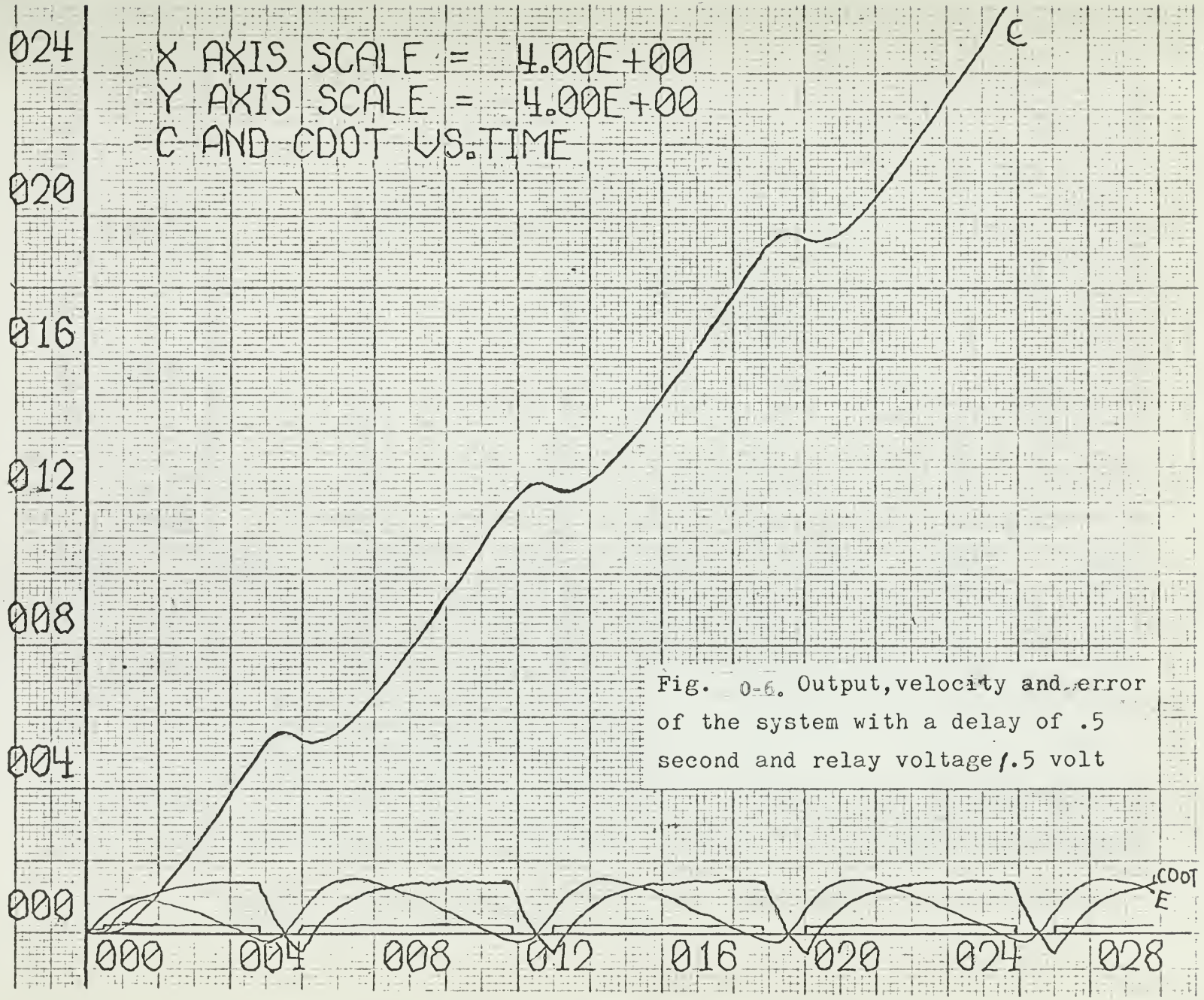


Fig. 0-6. Output, velocity and error of the system with a delay of .5 second and relay voltage .5 volt

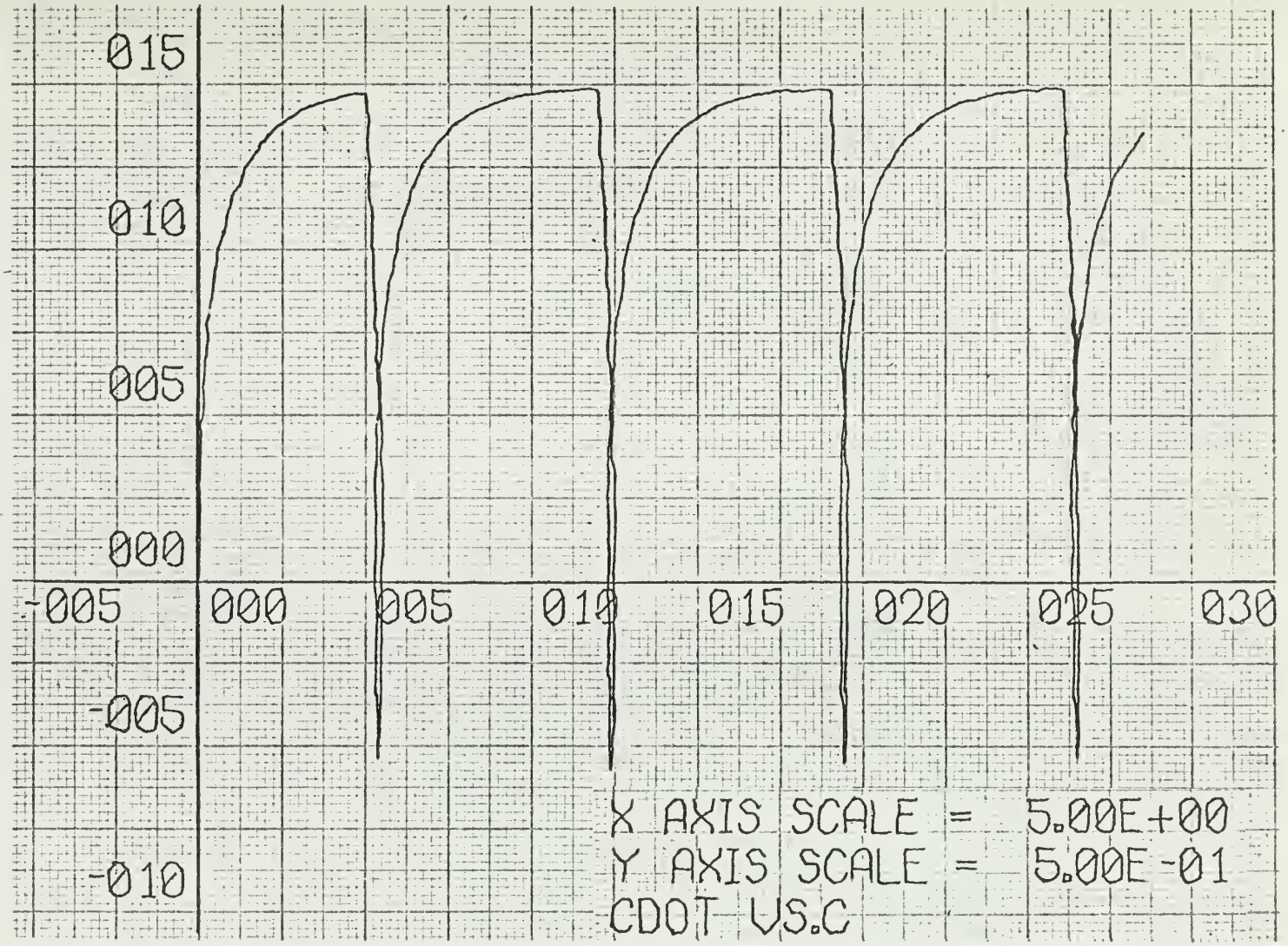


Fig 10-7. phase plane (rate output vs. output) of relay servo system with unit ramp input, time delay .5 second, k_t .02, relay voltage/.5 volt.

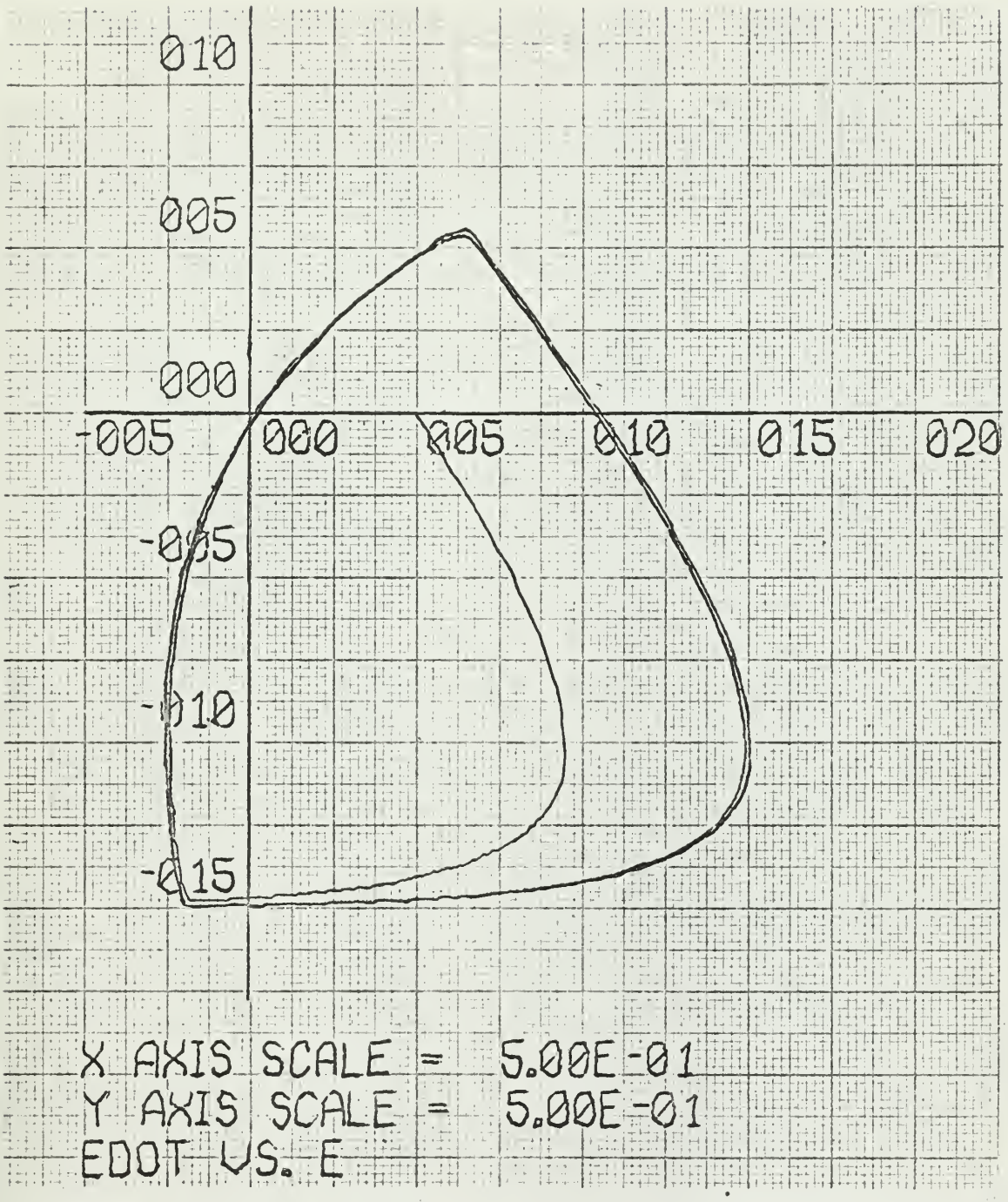


Fig. 10-8. Phase trajectory of ralay servo system to an unit ramp input, time delay = .5 second, relay voltage=/.5 volt.

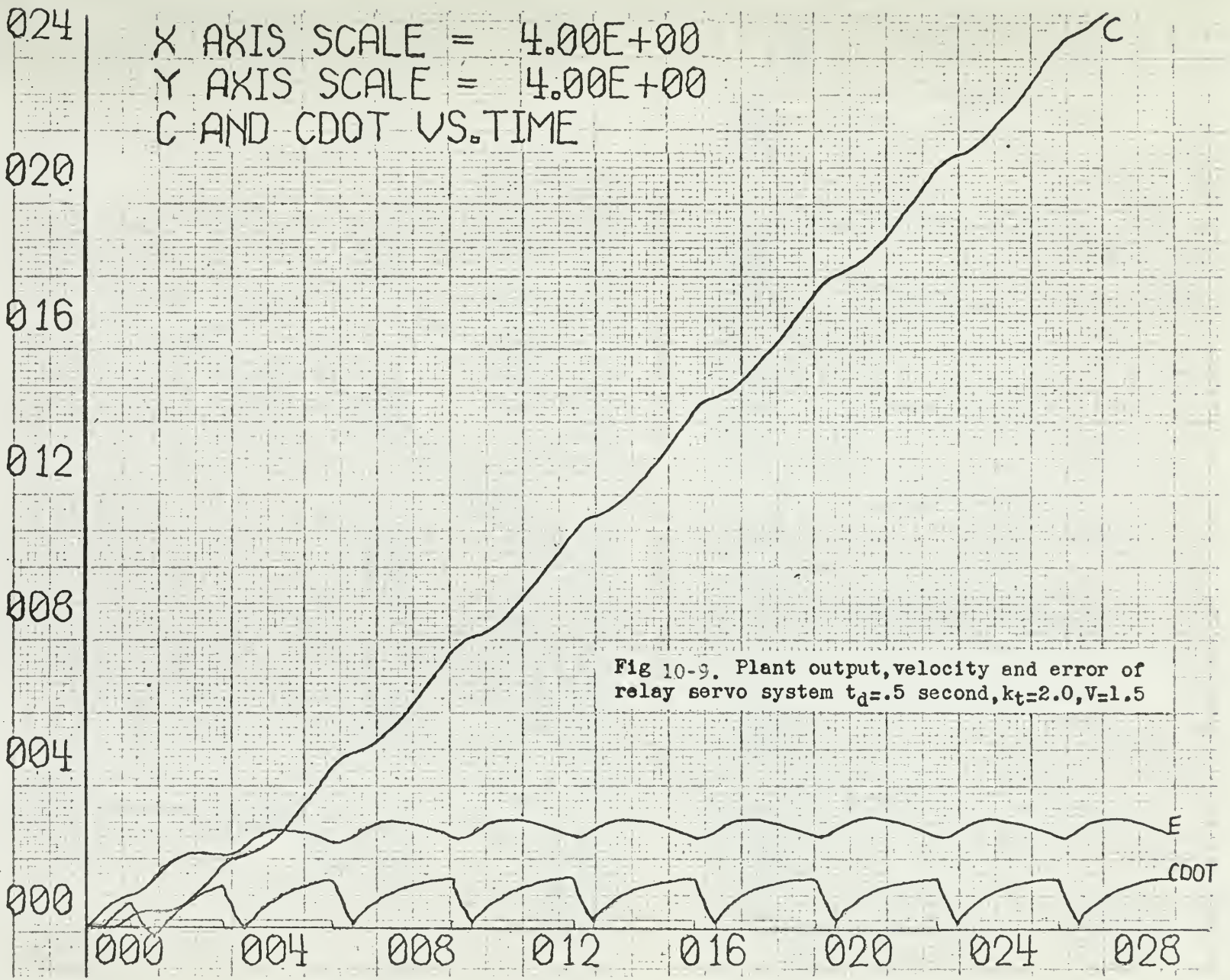


Fig 10-9. Plant output, velocity and error of relay servo system t_d=.5 second, k_t=2.0, V=1.5

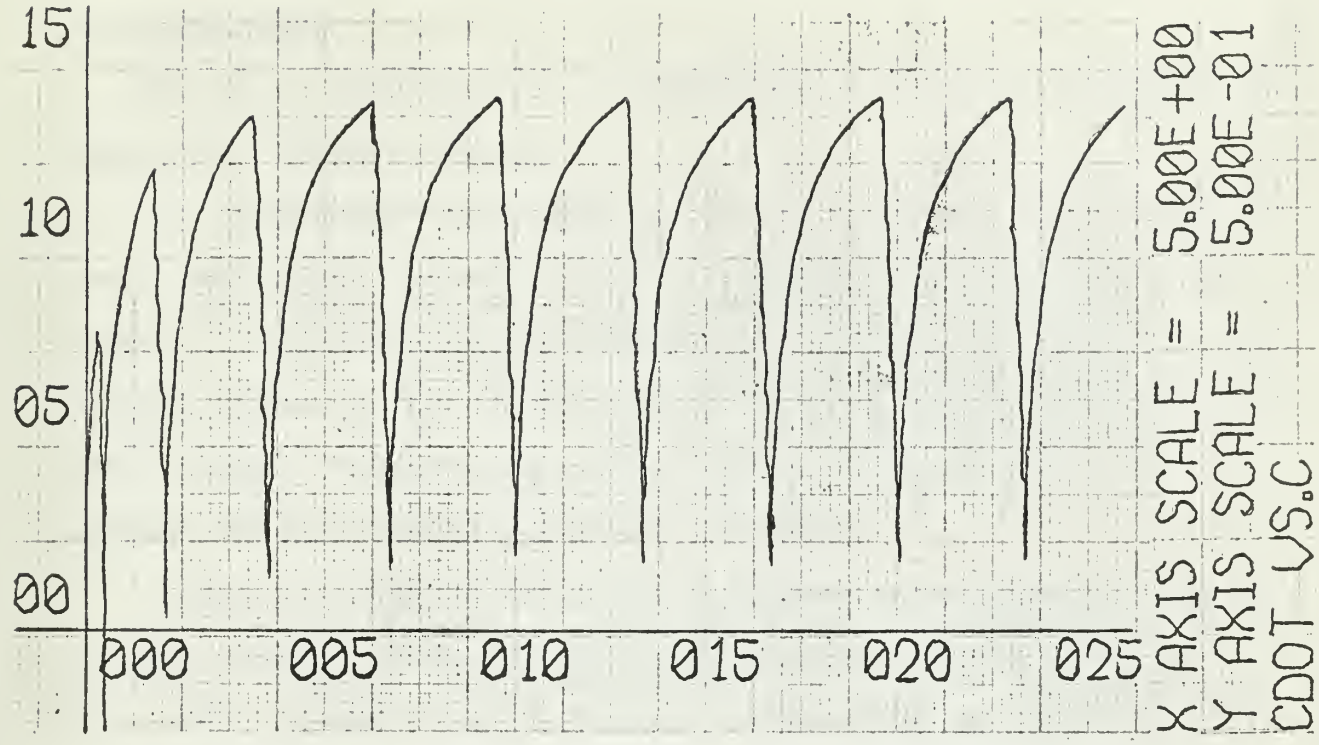


Fig. 10-10CDOT vs.C phase plane, t_d=.5 second, k_t=2.0, V=1.5.

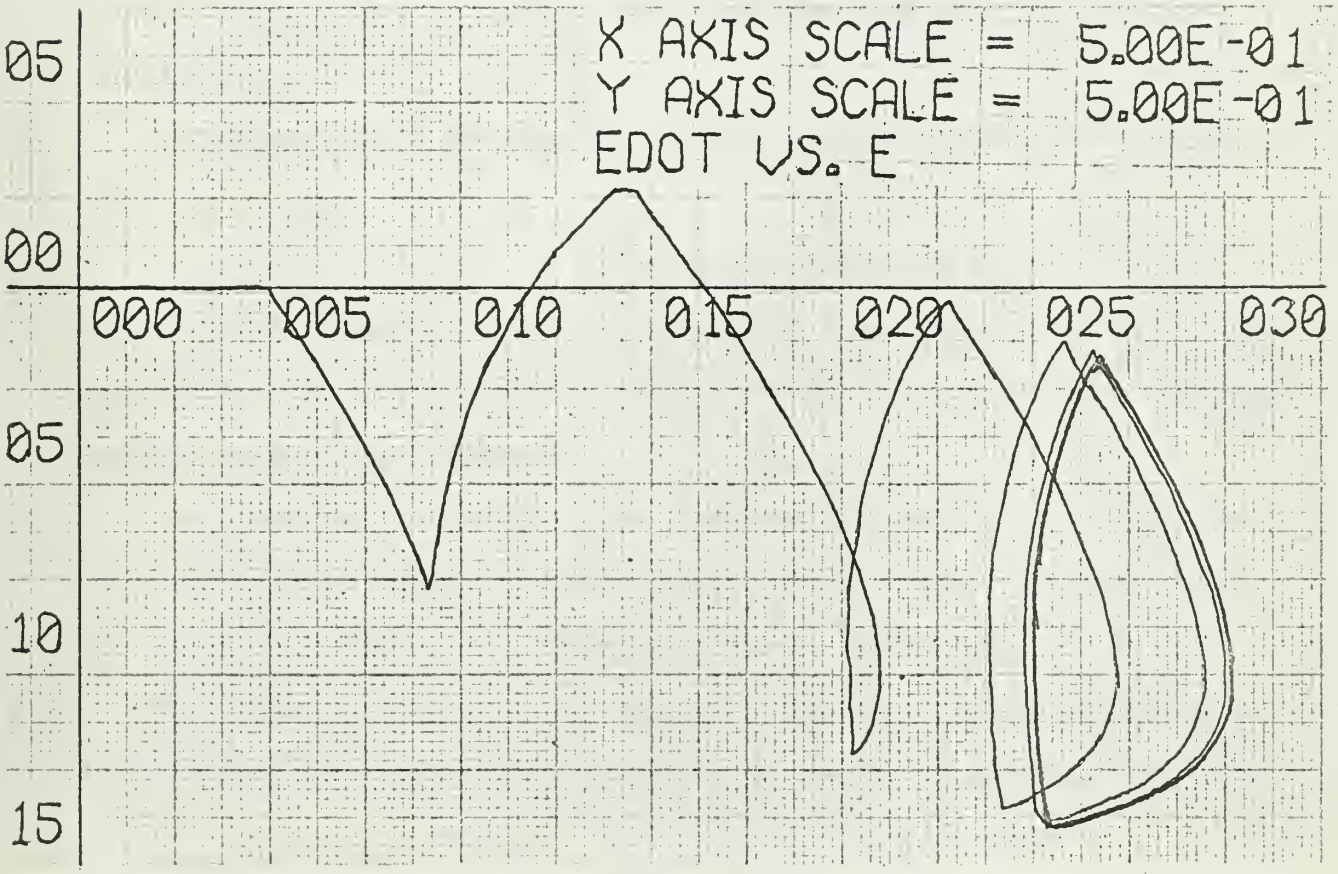


Fig 10-11. EDOT vs. E phase plane, t_d=.5 second, k_t=2.0, V=1.5.

11. Conclusions.

The aim of this thesis was to demonstrate the effects of transportation lag on the relay servo system.

For conducting the experiment by means of an analog computer, first one could simulate the time delay by Pade's approximation or by magnetic tape recorder. It was seen that the generation of time delay from magnetic tape recorder presents a more accurate and better result than that from the Pade's approximation. Therefore, the magnetic tape recorder was selected to generate the time delay throughout the experiment.

In this investigation most of the experiments were conducted by using a digital computer which gave more accurate results than the analog computer does; especially since it can give an exact value at any desired instant. The analog computer has also been used to check the results obtained from the digital computer. It is noticed that they are quite in agreement.

From the analysis and experiments the following conclusions has been obtained.

The root locus of the system was greatly changed due to the introduction of the transportation lag; and because time delay introduced periodicaly a phase shift of 360 degrees, it results in an additional (upper) branches when at high frequency.

The effect of time delay on the transient response is to increase the overshoot and undershoot and causes oscillations and also leads to a longer rise time, the response of the system can never be settled.

The effect of time delay on the phase plane is to rotate the switching line into the first and third quadrants, the results is growing oscillations

In other words the transportation lag will result in a terminal limit cycle which may or may not be acceptably small.

The plot of time delay versus oscillation amplitude came out to be a straight line; the delay time multiplied by a constant of .6 gives the magnitude of the oscillation amplitude.

The time delay versus steady state oscillation frequency when plotted on log log coordinates came out to be a straight line in such a relation that a constant of .135 divided by $t_d^{.57}$ gives the oscillation frequency.

The steady state error and the velocity are proportional to the time delay.

These effects caused by the transportation lag can be compensated to some extent. The method used in this investigation to eliminate steady state oscillation and to optimise the transient response by employing tachometer feedback. It was seen that a system with a delay time below .5 second can be compensated to an acceptable small oscillation amplitude (5 percent). But for time delay above .8 second it cannot be compensated to such a small amount; the only effect of increasing tachometer gain is to shift the whole response curve downward with the same amount of steady state over- and undershoot.

The system with a time delay below .5 second can be compensated to reach an acceptable optimum operation. For delay time below .5 second its relation with tachometer gain is approximately a straight line, therefore, it can be expressed by an equation.

When a time delay is present the relay type control can continuously oscillate, and with a magnitude of oscillation depending on the time delay. The relay type control can never be unstable in the sense that the amplitude of oscillation will continue to increase without limit. This is one advantage of the relay servo system.

BIBLIOGRAPHY

1. Donald P. Eckman: Automatic Process Control, John Wiley & Sons, New York, 1958.

2. W. J. Karplus and W. W. Soroka: Analog Methods, 2nd ed., McGraw-Hill Book Company, New York, 1959.

3. C. D. Morrill: A Sub-audio Time Delay Circuit, IRE Transaction on Electronic Computer, EC3 No.2, 1954, pp 45-49.

4. Yaohan Chu: Feedback Control Systems with Dead Time Lag or Distributed Lag by Root Locus Method, AIEE Transaction, Part 2, Industrial and Application, Vol. 71, Nov. 1952, pp 291-296.

5. Yaohan Chu: Synthesis of Feedback Control Systems by Phase Angle Loci, AIEE Transaction, Part 2, Industrial and Application, Vol. 71, Nov. 1952, pp 330-339.

6. G. J. Thaler and M. P. Pastel: Analysis and Design of Non-linear Feedback Control Systems, McGraw-Hill Book Company, New York, 1962.

7. Gorden J. Murphy: Control Engineering, Van Nostrand, New York, 1959.

8. D. Graham and D. McRuer: Analysis of Nonlinear Control Systems, John Wiley & Sons, New York, 1961.

9. David P. Donohue and Charles D. Federico: An Investigation of the Effects of Transportation Lag on Linear Feedback Control Systems, Master Thesis, The United States Naval Postgraduate School, 1961.

10. John C. West: Analytic Techniques for Nonlinear Control Systems, Van Nostrand, New York, 1960.

11. Otto J. Smith: Feedback Control Systems, McGraw-Hill Book Company, New York, 1958.

12. P. H. Hammond: Feedback Theory and Its Applications, Macmillan Company, New York, 1958.

13. John G. Truxal: Automativ Feedback Control System Synthesis, McGraw-Hill Book Company, 1955.